

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 19, 2004, 21:57:03 ; Search time 73.9837 Seconds  
(without alignments)  
504.271 Million cell updates/sec

Title: US-10-674-755-21

Perfect score: 518  
Sequence: 1 LAKKQTELEKLDNLDPGEK.....TKELDALNELGPDGDEB 104

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database :

A\_Geneseq\_23Sep04:\*  
1: geneseq1980s:\*  
2: geneseq1990s:\*  
3: geneseq2000s:\*  
4: geneseq2001s:\*  
5: geneseq2002s:\*  
6: geneseq2003as:\*  
7: geneseq2003bs:\*  
8: geneseq2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	508	98.1	213	7	ABW02601
2	508	98.1	8991	6	ABU08487
3	506	97.7	213	2	AAW14567
4	484	93.4	416	8	ADK52498
5	484	93.4	526	8	ADK52497
6	484	93.4	744	6	ABU00449
7	484	93.4	744	8	ADM92054
8	484	93.4	745	3	AAW1652
9	483	93.2	641	2	AAW61217
10	483	93.2	641	5	ABP54636
11	483	93.2	641	7	ADK45241
12	480	92.7	197	7	ABW02598
13	459	88.6	233	7	ABW02606
14	456.5	88.1	196	2	AAW14564
15	452	87.3	213	7	AAW14572
16	431	62.0	211	7	ABW02621
17	321	62.0	232	2	AAW14587
18	315	60.8	232	8	ADOS2055
19	315	60.8	232	8	ADOS2055
20	315	60.8	369	8	ADK52496
21	315	60.8	458	2	AAW14582
22	315	60.8	458	7	ABW02626
23	315	60.8	653	8	ADK52495
24	315	60.8	653	8	ADOS2080
25	306	59.1	212	2	AAW14588

26	306	59.1	212	7	ABW02622	Abw02622 Bg7817c p
27	304.5	58.8	233	2	AAW14590	Aaw14590 Streptoco
28	277	53.5	185	7	AD015316	Ad015316 S.pneumo
29	259	50.0	159	8	ABW02623	Abw02623 Bg7561c p
30	242.5	46.8	184	2	AAW14589	Aaw14589 Streptoco
31	199.5	38.5	487	8	ADR04321	Adr04321 Streptoco
32	199.5	38.5	489	8	ADOS2088	Ados2088 Streptoco
33	199.5	38.5	524	8	ADOS2082	Ados2082 E. coli B
34	199.5	38.5	627	8	ADOS2129	Ados2129 E. coli B
35	192.5	37.2	119	2	AAW46291	Aaw46291 Pneumono
36	192.5	37.2	215	7	AAW14563	Aaw14563 Streptoco
37	192.5	37.2	215	7	ABW02597	Abw02597 Atcc6303c
38	187.5	36.2	290	8	ADOS2119	Ados2119 pYA3637 b
39	187.5	36.2	298	8	ADOS2127	Ados2127 pYA3637 b
40	183	35.3	230	8	ADOS2086	Ados2086 S. pneumo
41	183	35.3	230	8	ADR04319	Adr04319 Streptoco
42	129.5	25.0	550	8	ADR48356	Adk48356 Streptoco
43	128	24.7	315	2	AAW04375	Aaw04375 Streptoco
44	128	24.7	619	2	AAW63437	Aaw63437 Pneumococ
45	128	24.7	619	2	AAW87598	Aaw87598 Pneumococ

## ALIGNMENTS

RESULT 1	ABW02601	standard; protein; 213 AA.
XX	ABW02601;	
AC	XX	
DT	12-FEB-2004	(first entry)
XX	XX	
DE	Bg8090c pneumococcal surface protein A (Pspa) central region.	
KM	Pneumococcal surface protein A; Pspa; diagnosis; antigenic; vaccine;	
XX	immunological; gene therapy; immunostimulant.	
OS	Unidentified.	
XX	XX	
FH	Key	Location/Qualifiers
FT	Misc-difference 2	/label= Unknown
XX	XX	
PN	US6592876-B1.	
PD	15-JUL-2003.	
XX	XX	
PF	15-SEP-1995;	95US-00529055.
PR	20-APR-1993;	93US-00048896.
PR	06-JUN-1995;	95US-00465746.
PA	(UABR-) UAB RES FOUND.	
XX	XX	
PI	Briles DE, McDaniel LS, Swiatlo E, Yocher J, Brooks-Walter A;	
XX	WPI; 2003-862841/80.	
DR	XX	
PT	Immunological composition for obtaining expression products used for	
PT	detecting the presence of Streptococcus pneumoniae or its strain,	
PT	completes at least two different full length isolated gene encoding	
XX	pneumococcal surface protein A.	
PS	Example 6; SEQ ID NO 47; 121pp; English.	
XX	XX	
CC	The present invention relates to an immunological composition comprising	
CC	at least 2 different full length isolated genes encoding pneumococcal	
CC	surface protein A (Pspa) from different groups based on restriction	
CC	fragment polymorphism analysis. The invention is useful for obtaining	
CC	expression products by recombinant techniques to detect, determine,	
CC	isolate or diagnose the presence of Streptococcus pneumoniae or its	
CC	strain. The expression product is useful for preparing antigenic,	

CC immunological or vaccine compositions, for eliciting antibodies, an  
 CC immunological response (other than or additional to antibodies) or a  
 CC protective response (including antibody or other immunological response  
 CC by administering compositions to a host). The invention is also useful as  
 CC vaccines and in gene therapy. The present sequence is Bg8090c  
 CC pneumococcal surface protein A (PspA) central region. This sequence is  
 CC used in the exemplification of the invention  
 XX  
 SQ Sequence 213 AA;  
 Query Match 98.1%; Score 508; DB 7; Length 213;  
 Best Local Similarity 98.1%; Pred. No. 1.5e-37;  
 Matches 102; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 LAKQTELEKLDNLDPGKTDDELDEKAFAEIDKKADELPNKVADLEKEISNLEILLG 60  
 DB 59 LAKQTELEKLDNLDPGKTDDELDEKAFAEIDKKADELPNKVADLEKEISNLEILLG 118  
 QY 61 GADPEDDTAALPNKLTATKKALEKTKOKELDPAALNELGPDGDEE 104  
 DB 119 GADPEDDTAALPNKLTATKKALEKTKOKELDPAALNELGPDGDEE 162  
 RESULT 2  
 ABU08487 standard; protein; 8991 AA.  
 XX  
 AC ABU08487;  
 DT 24-JUN-2003 (first entry)  
 XX  
 DE S. pneumoniae pneumococcal surface protein A (PspA) protein.  
 XX  
 KM Pneumococcal surface protein C; PspC; pneumococcal surface protein A;  
 KM alpha-helical; proline rich; repeat region; pneumococcal infection; PspA;  
 KM antibacterial.  
 XX  
 OS Streptococcus pneumoniae.  
 XX  
 FH Key Location/Qualifiers  
 FT Misc-difference 1..8991  
 FT /note="All Xaa residues within this sequence are  
 FT unknown"  
 XX  
 PN US6500613-B1.  
 PD 31-DEC-2002.  
 XX  
 PF 16-SEP-1996; 96US-00714741.  
 XX  
 PR 15-SEP-1995; 95US-00529055.  
 XX  
 PA (UTAL-) UNIV ALABAMA.  
 PI Briles DE, McDaniel LS, Swiatlo E, Yother J, Crain MJ;  
 PI Hollingshead S, Tart R, Brooks-Walter A;  
 XX  
 DR WPI; 2003-361534/34.  
 XX  
 PT Isolated PspC amino acid sequence used as polymerase chain reaction or  
 PT hybridization probe, comprises pneumococcal surface protein having alpha-  
 PT helical, proline rich and repeat regions.  
 XX  
 PS Disclosure; Col 145-188; 186pp; English.  
 XX  
 CC The present invention relates to the isolation of Streptococcus  
 CC pneumoniae pneumococcal surface protein C (PspC), and the polynucleotide  
 CC sequence encoding it. PspC is a pneumococcal surface protein A (PspA)-  
 CC like protein having alpha-helical, proline rich and repeat regions. The  
 CC PspC and PspA proteins may be used in a vaccine to protect against  
 CC pneumococcal infections. The polynucleotide sequences encoding PspC and  
 CC PspA may be used for the expression of the proteins, and as PCR primers  
 CC or hybridisation probes. The present sequence represents S. pneumoniae

CC PspA protein  
 XX  
 SQ Sequence 8991 AA;  
 Query Match 98.1%; Score 508; DB 6; Length 8991;  
 Best Local Similarity 98.1%; Pred. No. 1.1e-35;  
 Matches 102; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 LAKQTELEKLDNLDPGKTDDELDEKAFAEIDKKADELPNKVADLEKEISNLEILLG 60  
 DB 4625 LAKQTELEKLDNLDPGKTDDELDEKAFAEIDKKADELPNKVADLEKEISNLEILLG 4684  
 QY 61 GADPEDDTAALPNKLTATKKALEKTKOKELDPAALNELGPDGDEE 104  
 DB 4685 GADPEDDTAALPNKLTATKKALEKTKOKELDPAALNELGPDGDEE 4728  
 RESULT 3  
 AAW14567 standard; protein; 213 AA.  
 XX  
 AC AAW14567;  
 DT 17-OCT-2003 (revised)  
 DT 28-OCT-1997 (first entry)  
 XX  
 DE Streptococcus pneumoniae PspA central region.  
 XX  
 KM PspA; pneumococcal surface protein; vaccine; otitis media; meningitis;  
 KM bacteraemia; pneumonia.  
 XX  
 OS Streptococcus pneumoniae; strain Bg8090.  
 XX  
 FH Key Location/Qualifiers  
 FT Misc-difference 2 /note="unidentified amino acid"  
 FT  
 XX  
 PN WO9709994-A1.  
 PD 20-MAR-1997.  
 XX  
 PF 16-SEP-1996; 96WO-US014819.  
 XX  
 PR 15-SEP-1995; 95US-00529055.  
 XX  
 PA (UABR-) UAB RES FOUND.  
 PI Briles DE, McDaniel LS, Swiatlo E, Yother J, Crain MJ;  
 PI Hollingshead S, Tart R, Brooks-Walter A;  
 XX  
 DR WPI; 1997-202002/18.  
 XX  
 PT Streptococcus pneumoniae surface protein PspC and truncated PspA - used  
 PT in vaccines for protecting animals against S.pneumoniae infection.  
 XX  
 PS Example 6; Fig 13; 296pp; English.  
 XX  
 CC This sequence shows the central portion, including the C-terminus of the  
 CC alpha-helix region and some of the proline-rich region, of pneumococcal  
 CC surface protein A (PspA) of Streptococcus pneumoniae strain Bg8090.  
 CC Comparison of the N-terminal and central regions (AAW14533-57 and  
 CC AAW14562-91) of PspA polypeptides from different pneumococcal strains can  
 CC be used to divide the strains into several families based on sequence  
 CC homologies. PspA polypeptides, or fragments of them, can be used in  
 CC vaccines to protect animals against S. pneumoniae infection and hence for  
 CC the prevention of diseases such as otitis media, meningitis, bacteraemia  
 CC and pneumonia. The sequence of the 3' half of the PspA alpha-helical  
 CC region and the immediate 5' tip of the coding sequence are likely to be  
 CC the critical sequences for predicting PspA cross-reactions and vaccine  
 CC composition. (Updated on 17-OCT-2003 to standardise OS field)  
 XX  
 SQ Sequence 213 AA;

Query Match 97.7%; Score 506; DB 2; Length 213;  
 Best Local Similarity 98.1%; Pred. No. 2, 2e-37;  
 Matches 102; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 LAKKQTELEKLDNDPEGKTODELDKEAAEAELDKKADLPNKVADLEKEISNLEILLG 60  
 DB 59 LAKKQTELEKLDNDPEGKTODELDKEAAEAELDKKADLPNKVADLEKEISNLEILLG 118

QY 61 GADPEDDTAALPNKLTATKKAELKTKQKELDAALNELGPDGDEE 104  
 DB 119 GADPEDDTAALPNKLTATKKAELKTKQKELDAALNELGPDGDEE 162

RESULT 4  
 ADK52498  
 ID ADK52498 standard; protein; 416 AA.  
 AC ADK52498;  
 XX  
 XX  
 DT 20-MAY-2004 (first entry)  
 XX  
 XX alpha helical region PspA molecule from the EF3296 strain.  
 DE  
 XX Streptococcus pneumoniae infection; pneumococcal surface protein A; PspA;  
 KW hemolytic anemia disease; leukemia; lymphoma; sickle cell anemia;  
 KW Hodgkin's disease.  
 XX  
 OS Streptococcus pneumoniae.  
 XX  
 PN WO2004016231-A2.  
 XX  
 PD 26-FEB-2004.  
 XX  
 PF 17-FEB-2003; 2003WO-US008199.  
 XX  
 PR 15-MAR-2002; 2002US-0365351P.  
 XX  
 PA (UABR-) UAB RES FOUND.  
 XX  
 PI Briles DE;  
 XX  
 DR WPI; 2004-192068/18.  
 XX  
 PT Treating Streptococcus pneumoniae infection in a subject lacking a  
 functional spleen comprises administering an antibody that recognizes  
 PT pneumococcal surface protein A (PspA) or its binding portion.  
 XX  
 PS Claim 17; SEQ ID NO 4; 41pp; English.  
 XX  
 CC The present invention relates to treating Streptococcus pneumoniae  
 CC infection in a subject lacking a functional spleen comprises  
 CC administering an antibody that recognizes pneumococcal surface protein A  
 CC (PspA) or its binding portion. The method is useful for treating or  
 CC preventing Streptococcus pneumoniae infection in a subject lacking a  
 CC functional spleen. The disease-associated injury is especially due to  
 CC hemolytic anemia disease, leukemia or lymphoma, especially sickle cell  
 CC anemia or Hodgkin's disease. The present sequence represents the alpha  
 CC helical region PspA molecule from the EF3296 strain of Streptococcus  
 CC pneumoniae.  
 CC  
 SQ Sequence 416 AA;

Query Match 93.4%; Score 484; DB 8; Length 416;  
 Best Local Similarity 95.2%; Pred. No. 4, 5e-35;  
 Matches 99; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 LAKKQTELEKLDNDPEGKTODELDKEAAEAELDKKADLPNKVADLEKEISNLEILLG 60  
 DB 241 LAKKQTELEKLDNDPEGKTODELDKEAAEAELDKKADLPNKVADLEKEISNLEILLG 300

QY 61 GADPEDDTAALPNKLTATKKAELKTKQKELDAALNELGPDGDEE 104  
 DB 301 GADPEDDTAALPNKLTATKKAELKTKQKELDAALNELGPDGDEE 344

RESULT 5  
 ADK52497  
 ID ADK52497 standard; protein; 526 AA.  
 AC ADK52497;  
 XX  
 XX  
 DT 20-MAY-2004 (first entry)  
 XX  
 XX PspA molecule from the EF3296 strain of Streptococcus pneumoniae.  
 DE  
 XX Streptococcus pneumoniae infection; pneumococcal surface protein A; PspA;  
 KW hemolytic anemia disease; leukemia; lymphoma; sickle cell anemia;  
 KW Hodgkin's disease.  
 XX  
 OS Streptococcus pneumoniae.  
 XX  
 PN WO2004016231-A2.  
 XX  
 PD 26-FEB-2004.  
 XX  
 PF 17-FEB-2003; 2003WO-US008199.  
 XX  
 PR 15-MAR-2002; 2002US-0365351P.  
 XX  
 PA (UABR-) UAB RES FOUND.  
 XX  
 PI Briles DE;  
 XX  
 DR WPI; 2004-192068/18.  
 XX  
 PT Treating Streptococcus pneumoniae infection in a subject lacking a  
 functional spleen comprises administering an antibody that recognizes  
 PT pneumococcal surface protein A (PspA) or its binding portion.  
 XX  
 PS Claim 17; SEQ ID NO 3; 41pp; English.  
 XX  
 CC The present invention relates to treating Streptococcus pneumoniae  
 CC infection in a subject lacking a functional spleen comprises  
 CC administering an antibody that recognizes pneumococcal surface protein A  
 CC (PspA) or its binding portion. The method is useful for treating or  
 CC preventing Streptococcus pneumoniae infection in a subject lacking a  
 CC functional spleen. The disease-associated injury is especially due to  
 CC hemolytic anemia disease, leukemia or lymphoma, especially sickle cell  
 CC anemia or Hodgkin's disease. The present sequence represents PspA  
 CC molecule from the EF3296 strain of Streptococcus pneumoniae.  
 CC  
 SQ Sequence 526 AA;

Query Match 93.4%; Score 484; DB 8; Length 526;  
 Best Local Similarity 95.2%; Pred. No. 5, 9e-35;  
 Matches 99; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 LAKKQTELEKLDNDPEGKTODELDKEAAEAELDKKADLPNKVADLEKEISNLEILLG 60  
 DB 346 LAKKQTELEKLDNDPEGKTODELDKEAAEAELDKKADLPNKVADLEKEISNLEILLG 405

QY 61 GADPEDDTAALPNKLTATKKAELKTKQKELDAALNELGPDGDEE 104  
 DB 406 GADPEDDTAALPNKLTATKKAELKTKQKELDAALNELGPDGDEE 449

RESULT 6  
 ABU00449  
 ID ABU00449 standard; protein; 744 AA.  
 AC ABU00449;  
 XX  
 XX  
 DT 23-OCT-2003 (revised)  
 DT 11-FEB-2003 (first entry)  
 XX  
 XX S. pneumoniae type 4 strain protein from coding region #16.



KM antibacterial; antiinflammatory; meningitis; infection; diagnosis;  
 KM pneumococcal disease.  
 XX  
 OS Streptococcus pneumoniae.  
 XX  
 PN WO200006737-A2.  
 XX  
 PD 10-FEB-2000.  
 XX  
 PF 27-JUL-1999; 99WO-GB002451.  
 XX  
 PR 27-JUL-1998; 98GB-00016337.  
 PR 19-MAR-1999; 99US-0125164P.  
 XX  
 PA (MICK-) MICROBIAL TECHNIKS LTD.  
 XX  
 PI Gilbert CFG, Hanabro PM;  
 DR WPI; 2000-195300/17.  
 XX  
 PT New Streptococcal protein, useful as a vaccine, for diagnosis of  
 PT pneumococcal diseases and for screening agents capable of antagonizing or  
 PT inhibiting expression of the protein.  
 XX  
 PS Claim 2; Page 95; 108pp; English.  
 XX  
 CC AA01501 to AA01579 represent specifically claimed protein sequences  
 CC isolated from Streptococcus pneumoniae. AA015407 to AA015590 represent  
 CC specifically claimed nucleotide sequences isolated from S. pneumoniae.  
 CC The sequences have antibacterial and antiinflammatory properties. The  
 CC protein sequences, and fragments of them, are useful as immunogens and/or  
 CC antigens. The nucleotide sequences can be used in vaccines and in  
 CC diagnostic assays. The proteins and nucleotides can be useful for the  
 CC detection and diagnosis of S. pneumoniae. The protein sequences are also  
 CC useful for screening an agent capable of antagonizing, inhibiting or  
 CC interfering with the function or expression of the proteins in which the  
 CC agent is useful for treatment or prophylaxis of S. pneumoniae infection  
 CC and meningitis. AA015591 to AA015614 represent primers used in the  
 CC exemplification of the present invention  
 CC  
 XX  
 SQ Sequence 745 AA;  
 Query Match 93.4%; Score 484; DB 3; Length 745;  
 Best Local Similarity 95.2%; Pred. No. 8.8e-35;  
 Matches 99; Conservative 1; Mismatches 4; Indels 0; Gaps 0;  
 QY 1 LAKKQTELEKLDNLDPEKTDDELDEKAAEAELDKKADLPNKVADLEKEISNLEILLG 60  
 DB 346 LAKKQTELEKLDNLDPEKTDDELDEKAAEAELDKKADLPNKVADLEKEISNLEILLG 405  
 QY 61 GADPEDDTAALPNKLTATKKAELKTKQKELDAALNLELGPDDDEE 104  
 DB 406 GADPEDDTAALONKLTATKKAELKTKQKELDAALNLELGPDDDEE 449  
 RESULT 9  
 AA061217  
 ID AA061217 standard; protein; 641 AA.  
 XX  
 AC AA061217;  
 XX  
 DT 02-OCT-1998 (first entry)  
 XX  
 DB Streptococcus pneumoniae SP0092 protein.  
 XX  
 KW Streptococcus pneumoniae; antigen; vaccine; infection; diagnosis;  
 KM detection; pneumonia; otitis media; meningitis.  
 XX  
 OS Streptococcus pneumoniae.  
 XX  
 FH Key Location/Qualifiers  
 FT Misc-difference 306  
 FT /label= unknown

FT /note= "encoded by NCR"  
 XX  
 PN WO9818930-A2.  
 XX  
 PD 07-MAY-1998.  
 XX  
 PF 30-OCT-1997; 97WO-US019422.  
 XX  
 PR 31-OCT-1996; 96US-0029960P.  
 XX  
 PA (HUMA-) HUMAN GENOME SCI INC.  
 XX  
 PI Kunesh CA, Choi GH, Johnson LS, Hromockyj A;  
 PI  
 DR WPI; 1998-272224/24.  
 DR N-PSDB; AAV27403.  
 XX  
 PT Nucleic acid encoding antigenic peptide(s) from Streptococcus pneumoniae  
 PT - or their epitope-containing fragments, useful in protective or  
 PT therapeutic vaccines, and for diagnosis.  
 XX  
 PS Claim 11; Page 82; 118pp; English.  
 XX  
 CC The present sequence represents a protein from Streptococcus pneumoniae.  
 CC The nucleic acid sequence encoding the Streptococcus pneumoniae protein  
 CC can be useful in vaccines for inducing protective antibodies against  
 CC Streptococcus pneumoniae, for treatment or prevention of infection e.g.  
 CC pneumonia, otitis media or meningitis. Probes based on the nucleic acid  
 CC are used to detect Streptococcus infection (by usual hybridisation or  
 CC amplification methods), also for isolating Streptococcus genes or their  
 CC allelic variants. The protein can be used similarly to detect specific  
 CC antibodies in standard immunoassays, especially for diagnosing or  
 CC monitoring infections. Antibodies which bind the protein are used to  
 CC detect corresponding antigens, to purify the protein and for passive  
 CC immunisation (optionally coupled to a toxin). Vaccines are administered,  
 CC e.g. by injection, orally or through the skin, typically at 0.01-1000  
 CC (especially 10-300) mu g/ml per dose  
 CC  
 XX  
 SQ Sequence 641 AA;  
 Query Match 93.2%; Score 483; DB 2; Length 641;  
 Best Local Similarity 95.2%; Pred. No. 9e-35;  
 Matches 99; Conservative 1; Mismatches 4; Indels 0; Gaps 0;  
 QY 1 LAKKQTELEKLDNLDPEKTDDELDEKAAEAELDKKADLPNKVADLEKEISNLEILLG 60  
 DB 243 LAKKQTELEKLDNLDPEKTDDELDEKAAEAELDKKADLPNKVADLEKEISNLEILLG 302  
 QY 61 GADPEDDTAALPNKLTATKKAELKTKQKELDAALNLELGPDDDEE 104  
 DB 303 GADPEDDTAALONKLTATKKAELKTKQKELDAALNLELGPDDDEE 346  
 RESULT 10  
 ABP54636  
 ID ABP54636 standard; protein; 641 AA.  
 XX  
 AC ABP54636;  
 XX  
 DT 04-SEP-2002 (first entry)  
 XX  
 DB S. pneumoniae SP092 protein sequence SEQ ID NO:160.  
 XX  
 KW Streptococcus pneumoniae; epitope; vaccine; antigenic protein;  
 KM antibacterial; Streptococcal infection; detection.  
 XX  
 OS Streptococcus pneumoniae.  
 XX  
 PN US2002061545-A1.  
 XX  
 PD 23-MAY-2002.  
 XX  
 PF 22-JAN-2001; 2001US-00765272.

XX 30-OCT-1997; 97US-00961083.  
PR (CHOI/) CHOI G H.  
XX (KUNS/) KUNSCH C A.  
PA (BARA/) BARASH S C.  
PA (DILL/) DILLON P J.  
PA (DOUG/) DOUGHERTY B.  
PA (FANN/) FANNON M R.  
PA (ROSE/) ROSEN C A.  
XX  
PI Choi GH, Kunsch CA, Barash SC, Dillon PJ, Dougherty B, Fannon MR,  
PI Rosen CA;  
XX WPI: 2003-764574/72.  
XX N-PSDB; ADC45240.  
XX  
XX New Streptococcus pneumoniae antigens, useful for detecting Streptococcus  
PT and for preventing or attenuating disease caused by Streptococcus  
PT infection.  
XX  
XX Claim 11; Page 43; 70pp; English.  
XX  
XX ABQ84792 to ABQ84904 represents nucleic acids which encode the  
CC Streptococcus pneumoniae antigens given in ABP54557 to ABP54669. The S.  
CC pneumoniae antigens have antibacterial activity and can be used in  
CC vaccines. The S. pneumoniae antigens can also be used to prevent or  
CC attenuate a Streptococcus pneumoniae infection in an animal. The polynucleotides  
CC encoding the S. pneumoniae antigens can be used to detect Streptococcus  
CC nucleic acids. ABQ84905 to ABQ85130 represent primers used in the cloning  
CC of S. pneumoniae ORFs (open reading frames) which are used in an example  
CC from the present invention  
XX  
SQ Sequence 641 AA;

Query Match 93.2%; Score 483; DB 5; Length 641;  
Best Local Similarity 95.2%; Pred. No. 9e-35;  
Matches 99; Conservative 1; Mismatches 4; Indels 0; Gaps 0;  
QY 1 LAKKQTELEKLDNLDPBGKTODELDKEAAEALDKKADDELPNKVADLEKEISNLEITLLG 60  
DB 243 LAKKQTELEKLDNLDPBGKTODELDKEAAEALDKKADDELQNVADLEKEISNLEITLLG 302  
QY 61 GADPEDDTAALPNKLTATKKALEKTKQKELDAALNELGPDGDDEE 104  
DB 303 GADKEDDTAALQNKLTATKKALEKTKQKELDAALNELGPDGDDEE 346

RESULT 11  
ADC45241  
ID ADC45241 standard; protein; 641 AA.  
XX  
AC ADC45241;  
XX  
DT 18-DEC-2003 (first entry)  
XX  
DE S. pneumoniae antigenic protein SP092.  
XX  
KM Antigen; bacterial infection; vaccine; pneumonia; antibacterial.  
XX  
OS Streptococcus pneumoniae.  
XX  
PN US6573082-B1.  
XX  
PD 03-JUN-2003.  
XX  
PF 28-MAR-2000; 2000US-00536784.  
XX  
PR 31-OCT-1996; 96US-0029960P.  
PR 30-OCT-1997; 97US-00961083.  
XX  
PA (HUMA-) HUMAN GENOME SCI INC.  
XX

PI Choi GH, Kunsch CA, Barash SC, Dillon PJ, Dougherty B, Fannon MR;  
PI Rosen CA;  
XX WPI: 2003-764574/72.  
XX N-PSDB; ADC45240.  
XX  
XX Novel polynucleotide encoding Streptococcus pneumoniae polypeptides  
PT useful for producing vaccines for prevention or attenuation of infection  
PT by Streptococcus pneumoniae.  
XX  
PS Example 1; SEQ ID NO 160; 58pp; English.  
XX  
XX The invention relates to an isolated polynucleotide consisting of a  
CC Streptococcus pneumoniae nucleic acid (appearing as ADC45122 and encoding  
CC SP028) one of 113 disclosed nucleic acids encoding 113 S. pneumoniae  
CC antigens. Also included are making a recombinant vector by inserting the  
CC nucleic acid into a vector, an isolated polynucleotide consisting of at  
CC least 50 or 100 contiguous nucleotides of the SP028 nucleic acid, and a  
CC recombinant host cell comprising the SP028 polynucleotide. The nucleic  
CC acids are useful as DNA vaccine against Streptococcus pneumoniae  
CC infection (e.g. pneumonia). Nucleic acids derived from the S. pneumoniae  
CC antigen nucleic acids are useful as probes for use in diagnostic methods  
CC for detecting S. pneumoniae gene expression. The present sequence  
CC represents an S. pneumoniae antigenic protein.  
XX  
SQ Sequence 641 AA;

Query Match 93.2%; Score 483; DB 7; Length 641;  
Best Local Similarity 95.2%; Pred. No. 9e-35;  
Matches 99; Conservative 1; Mismatches 4; Indels 0; Gaps 0;  
QY 1 LAKKQTELEKLDNLDPBGKTODELDKEAAEALDKKADDELPNKVADLEKEISNLEITLLG 60  
DB 243 LAKKQTELEKLDNLDPBGKTODELDKEAAEALDKKADDELQNVADLEKEISNLEITLLG 302  
QY 61 GADPEDDTAALPNKLTATKKALEKTKQKELDAALNELGPDGDDEE 104  
DB 303 GADKEDDTAALQNKLTATKKALEKTKQKELDAALNELGPDGDDEE 346

RESULT 12  
ABW02598  
ID ABW02598 standard; protein; 197 AA.  
XX  
AC ABW02598;  
XX  
DT 12-FEB-2004 (first entry)  
XX  
DE Act122c pneumococcal surface protein A (PspA) central region.  
XX  
KM Pneumococcal surface protein A; PspA; diagnosis; antigenic; vaccine;  
KM immunological; gene therapy; immunostimulant.  
XX  
OS Unidentified.  
XX  
PN US6592876-B1.  
XX  
PD 15-JUL-2003.  
XX  
PF 15-SEP-1995; 95US-00529055.  
XX  
PR 20-APR-1993; 93US-00048896.  
PR 06-JUN-1995; 95US-00465746.  
XX  
PA (UABR-) UAB RES FOUND.  
XX  
PI Briles DR, McDaniel LS, Swiatlo E, Yother J, Brooks-Walter A;  
XX WPI: 2003-862841/80.  
XX  
XX Immunological composition for obtaining expression products used for  
PT detecting the presence of Streptococcus pneumoniae or its strain,  
PT comprises at least two different full length isolated gene encoding

PT pneumococcal surface protein A.  
 XX  
 PS Example 6; SEQ ID NO 44; 121bp; English.  
 CC The present invention relates to an immunological composition comprising  
 CC at least 2 different full length isolated genes encoding pneumococcal  
 CC surface protein A (PspA) from different groups based on restriction  
 CC fragment polymorphism analysis. The invention is useful for obtaining  
 CC expression products by recombinant techniques to detect, determine,  
 CC isolate or diagnose the presence of Streptococcus pneumoniae or its  
 CC strain. The expression product is useful for preparing antigenic,  
 CC immunological or vaccine compositions, for eliciting antibodies, an  
 CC immunological response (other than or additional to antibodies) or a  
 CC protective response (including antibody or other immunological response  
 CC by administering compositions to a host). The invention is also useful as  
 CC vaccine and in gene therapy. The present sequence is Ac122c pneumococcal  
 CC surface protein A (PspA) central region. This sequence is used in the  
 CC exemplification of the invention  
 XX  
 SQ Sequence 197 AA;

Query Match 92.7%; Score 480; DB 7; Length 197;  
 Best Local Similarity 93.3%; Pred. No. 4.3e-35;  
 Matches 97; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 LAKKOTELKLDNDPBGKTODELDKEAAEALDKKADLPKVADEKEISNLEILLG 60  
 DB 22 LAOKPTGLEKLDSDPBGKTODELDKEAGEALDKKADGLPKVSDLEKEISNLEILLG 81  
 QY 61 GADPEDDTAALPNKLTATKKALEKTKQKELDPAALNLEIGPDGDEEE 104  
 DB 82 GADSEDDTAALPNKLTATKKALEKTKQKELDPAALNLEIGPDGDEEE 125

RESULT 13  
 ABM02606  
 ID ABM02606 standard; protein; 233 AA.  
 XX  
 AC ABM02606;

DT 12-FEB-2004 (first entry)  
 XX  
 DE Efl109c pneumococcal surface protein A (PspA) central region.  
 XX  
 KM Pneumococcal surface protein A; PspA; diagnosis; antigenic; vaccine;  
 KM immunological; gene therapy; immunostimulant.  
 XX  
 OS Unidentified.

PH Key Location/Qualifiers  
 FT Misc-difference 1..233  
 FT /note="Xaa = Unknown amino acid"

US6592876-B1.  
 XX  
 PN 15-JUL-2003.  
 PD  
 XX 15-SEP-1995; 95US-00529055.  
 PF  
 XX 20-APR-1993; 93US-00048896.  
 PR 06-JUN-1995; 95US-00465746.  
 XX  
 PA (UABR-) UAB RES FOUND.  
 PI Briles DE, McDaniel LS, Swiatlo E, Yocher J, Brooke-Walter A;  
 XX  
 DR WPI; 2003-862841/80.

PT Immunological composition for obtaining expression products used for  
 PT detecting the presence of Streptococcus pneumoniae or its strain,  
 PT comprising at least two different full length isolated gene encoding  
 PT pneumococcal surface protein A.  
 XX

PS Example 6; SEQ ID NO 52; 121bp; English.  
 XX  
 CC The present invention relates to an immunological composition comprising  
 CC at least 2 different full length isolated genes encoding pneumococcal  
 CC surface protein A (PspA) from different groups based on restriction  
 CC fragment polymorphism analysis. The invention is useful for obtaining  
 CC expression products by recombinant techniques to detect, determine,  
 CC isolate or diagnose the presence of Streptococcus pneumoniae or its  
 CC strain. The expression product is useful for preparing antigenic,  
 CC immunological or vaccine compositions, for eliciting antibodies, an  
 CC immunological response (other than or additional to antibodies) or a  
 CC protective response (including antibody or other immunological response  
 CC by administering compositions to a host). The invention is also useful as  
 CC vaccine and in gene therapy. The present sequence is Efl109c  
 CC pneumococcal surface protein A (PspA) central region. This sequence is  
 CC used in the exemplification of the invention  
 XX  
 SQ Sequence 233 AA;

Query Match 88.6%; Score 459; DB 7; Length 233;  
 Best Local Similarity 90.4%; Pred. No. 3.9e-33;  
 Matches 94; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

QY 1 LAKKOTELKLDNDPBGKTODELDKEAAEALDKKADLPKVADEKEISNLEILLG 60  
 DB 53 LAOKPTGLEKLDSDPBGKTODELDKEAGEALDKKADLPKVADEKEISNLEILLG 112  
 QY 61 GADPEDDTAALPNKLTATKKALEKTKQKELDPAALNLEIGPDGDEEE 104  
 DB 113 GADSEDDTAALPNKLTATKKALEKTKQKELDPAALNLEIGPDGDEEE 156

RESULT 14  
 AAM14564  
 ID AAM14564 standard; protein; 196 AA.  
 XX  
 AC AAM14564;

DT 17-OCT-2003 (revised)  
 DT 28-OCT-1997 (first entry)  
 XX

DE Streptococcus pneumoniae PspA central region.  
 XX  
 KM PspA; pneumococcal surface protein; vaccine; otitis media; meningitis;  
 KM bacteraemia; pneumonia.  
 XX  
 OS Streptococcus pneumoniae; strain Ac122.

PN WO9709994-A1.  
 XX  
 PD 20-MAR-1997.  
 XX  
 PF 16-SEP-1996; 96WO-US014819.  
 XX  
 PR 15-SEP-1995; 95US-00529055.  
 XX

PA (UABR-) UAB RES FOUND.  
 PI Briles DE, McDaniel LS, Swiatlo E, Yocher J, Crain MJ;  
 PI Hollingshead S, Tart R, Brooks-Walter A;  
 XX  
 DR WPI; 1997-202002/18.

PT Streptococcus pneumoniae surface protein PspC and truncated PspA - used  
 PT in vaccines for protecting animals against S.pneumoniae infection.  
 XX

PS Example 6; Fig 13; 296bp; English.

XX This sequence shows the central portion, including the C-terminus of the  
 CC alpha-helix region and some of the proline-rich region, of pneumococcal  
 CC surface protein A (PspA) of Streptococcus pneumoniae strain Ac122.  
 CC Comparison of the N-terminal and central regions (AAM14533-57 and  
 CC AAM14562-91) of PspA polypeptides from different pneumococcal strains can

CC be used to divide the strains into several families based on sequence  
 CC homologies. PspA polypeptides, or fragments of them, can be used in  
 CC vaccines to protect animals against S. pneumoniae infection and hence for  
 CC the prevention of diseases such as otitis media, meningitis, bacteraemia  
 CC and pneumonia. The sequence of the 3' half of the PspA alpha-helical  
 CC region and the immediate 5' tip of the coding sequence are likely to be  
 CC the critical sequences for predicting PspA cross-reactions and vaccine  
 CC composition. (Updated on 17-OCT-2003 to standardise OS field)

XX Sequence 196 AA;

Query Match 88.1%; Score 456.5; DB 2; Length 196;  
 Best Local Similarity 91.3%; Pred. No. 5.4e-33;  
 Matches 95; Conservative 3; Mismatches 5; Indels 1; Gaps 1;

QY 1 LAKKQTELEKLDNLDPEGKTODELDKEAAEALDKKADLPNKVADLEKEISNLEITLG 60  
 ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
 Db 22 LAQKPTGLEKLDLSDPGKRYQDELDPKEAGEALDKKADLPNKVADLEKEISNLEITLG 81

QY 61 GADPEDDTAALPNKLTATYKAELEKTOKELDPAALNELGPDGDEER 104  
 |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
 Db 82 GADSEDDTAALPNKLTATYKAELEKTKELDPAALNELGPDGDEER 124

RESULT 15  
 AAW14572  
 ID AAW14572 standard; protein; 233 AA.  
 XX

AC AAW14572;  
 XX  
 DT 17-OCT-2003 (revised)  
 DT 28-OCT-1997 (first entry)  
 XX

DE Streptococcus pneumoniae PspA central region.

KM PspA; pneumococcal surface protein; vaccine; otitis media; meningitis;  
 KW bacteraemia; pneumonia.  
 XX

OS Streptococcus pneumoniae; strain Efs3296.

XX Key Location/Qualifiers  
 FH Misc-difference 129 /note= "unidentified amino acid"  
 FT Misc-difference 131  
 FT Misc-difference 131 /note= "unidentified amino acid"  
 XX

PN W09709994-A1.

PD 20-MAR-1997.

PF 16-SEP-1996; 96WO-US014819.

PR 15-SEP-1995; 95US-00529055.

PA (UABR-) UAB RES FOUND.

PI Briles DE, McDaniel LS, Swiatlo E, Yother J, Crain MJ;  
 PI Hollingshead S, Tart R, Brooks-Walter A;

XX WPI; 1997-202002/18.

PT Streptococcus pneumoniae surface protein PspC and truncated PspA - used  
 PT in vaccines for protecting animals against S.pneumoniae infection.

XX Example 6; Fig 13; 296pp; English.

CC This sequence shows the central portion, including the C-terminus of the  
 CC alpha-helix region and some of the proline-rich region, of pneumococcal  
 CC surface protein A (PspA) of Streptococcus pneumoniae strain Efs3296.  
 CC Comparison of the N-terminal and central regions (AAW14533-57 and  
 CC AAW14562-91) of PspA polypeptides from different pneumococcal strains can  
 CC be used to divide the strains into several families based on sequence  
 CC homologies. PspA polypeptides, or fragments of them, can be used in

CC vaccines to protect animals against S. pneumoniae infection and hence for  
 CC the prevention of diseases such as otitis media, meningitis, bacteraemia  
 CC and pneumonia. The sequence of the 3' half of the PspA alpha-helical  
 CC region and the immediate 5' tip of the coding sequence are likely to be  
 CC the critical sequences for predicting PspA cross-reactions and vaccine  
 CC composition. (Updated on 17-OCT-2003 to standardise OS field)

XX Sequence 233 AA;

Query Match 87.3%; Score 452; DB 2; Length 233;  
 Best Local Similarity 89.4%; Pred. No. 1.7e-32;  
 Matches 93; Conservative 2; Mismatches 9; Indels 0; Gaps 0;

QY 1 LAKKQTELEKLDNLDPEGKTODELDKEAAEALDKKADLPNKVADLEKEISNLEITLG 60  
 ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
 Db 53 LAQKPTGLEKLDLSDPGKRYQDELDPKEAGEALDKKADLPNKVADLEKEISNLEITLG 112

QY 61 GADPEDDTAALPNKLTATYKAELEKTOKELDPAALNELGPDGDEER 104  
 |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
 Db 113 GADSEDDTAALPNKLTATYKAELEKTKELDPAALNELGPDGDEER 156

Search completed: November 19, 2004, 22:20:18  
 Job time : 74.9837 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 19, 2004, 22:06:04 ; Search time 18.813 seconds  
(without alignments)  
366.612 Million cell updates/sec

Title: US-10-674-755-21

Perfect score: 518  
Sequence: 1 LAKKQTELEKLDNDPEKGT.....TQKELDAALNELGPDDEEE 104

Scoring table: BLOSUM62

Gapop 10.0 , Gapept 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:\*  
1: /cgn2\_6/prodata/1/aa/5A COMB.pep:\*  
2: /cgn2\_6/prodata/1/aa/5B COMB.pep:\*  
3: /cgn2\_6/prodata/1/aa/6A COMB.pep:\*  
4: /cgn2\_6/prodata/1/aa/6B COMB.pep:\*  
5: /cgn2\_6/prodata/1/aa/6C COMB.pep:\*  
6: /cgn2\_6/prodata/1/aa/backfile1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	518	100.0	104	4	US-09-147-875A-21
2	512	98.8	104	2	US-08-710-749-20
3	508	98.1	213	4	US-08-529-055-47
4	508	98.1	8991	4	US-08-714-741-32
5	500	96.5	104	2	US-08-710-749-19
6	500	96.5	104	4	US-09-147-875A-20
7	483	93.2	641	3	US-08-961-083-160
8	483	93.2	641	4	US-09-536-784-160
9	480	92.7	197	4	US-08-529-055-44
10	479	92.5	102	2	US-08-710-749-21
11	479	92.5	102	4	US-09-147-875A-18
12	459	88.6	233	4	US-08-529-055-52
13	375	72.4	80	2	US-09-147-875A-19
14	375	72.4	80	4	US-08-710-749-18
15	321	62.0	108	2	US-08-710-749-26
16	321	62.0	108	4	US-09-147-875A-23
17	321	62.0	211	4	US-08-529-055-67
18	315	60.8	108	2	US-08-710-749-24
19	315	60.8	108	4	US-09-147-875A-25
20	315	60.8	232	4	US-08-529-055-70
21	315	60.8	458	4	US-08-529-055-73
22	313	60.4	108	4	US-09-147-875A-24
23	311	60.0	106	4	US-09-147-875A-22
24	306	59.1	212	4	US-08-529-055-68
25	303	58.5	108	2	US-08-710-749-22
26	303	58.5	108	4	US-08-710-749-23
27	266	51.4	108	2	US-08-710-749-25

28	266	51.4	108	4	US-09-147-875A-26	Sequence 26, Appl
29	259	50.0	185	4	US-08-529-055-69	Sequence 69, Appl
30	192.5	37.2	119	2	US-08-710-749-27	Sequence 27, Appl
31	192.5	37.2	119	4	US-09-147-875A-27	Sequence 27, Appl
32	192.5	37.2	215	4	US-08-529-055-43	Sequence 43, Appl
33	129.5	25.0	550	4	US-09-583-110-4871	Sequence 4871, Ap
34	128	24.7	288	3	US-08-312-949-4	Sequence 4, Appl
35	128	24.7	288	3	US-08-446-201-4	Sequence 4, Appl
36	128	24.7	619	1	US-08-465-746-2	Sequence 2, Appl
37	128	24.7	619	1	US-08-214-164-2	Sequence 2, Appl
38	128	24.7	619	2	US-08-467-852A-3	Sequence 3, Appl
39	128	24.7	619	2	US-08-246-636-2	Sequence 2, Appl
40	128	24.7	619	2	US-08-247-491A-3	Sequence 3, Appl
41	128	24.7	619	2	US-08-319-795-2	Sequence 2, Appl
42	128	24.7	619	2	US-08-468-985-2	Sequence 2, Appl
43	128	24.7	619	3	US-08-312-949-2	Sequence 2, Appl
44	128	24.7	648	1	US-08-072-070-2	Sequence 2, Appl
45	128	24.7	648	1	US-08-469-434-2	Sequence 2, Appl

## ALIGNMENTS

```

RESULT 1
US-09-147-875A-21
; Sequence 21, Application US/09147875A
; Patent No. 6638516
; GENERAL INFORMATION:
; APPLICANT: BECKER et al.
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS
; FILE REFERENCE: 454312-2471
; CURRENT APPLICATION NUMBER: US/09/147,875A
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 21
; LENGTH: 104
; TYPE: PRP
; ORGANISM: Streptococcus pneumoniae
US-09-147-875A-21

Query Match      100.0%; Score 518; DB 4; Length 104;
Best Local Similarity 100.0%; Pred. No. 6,9e-43;
Matches 104; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 LAKKQTELEKLDNDPEKGTODELDKEAAEALDKKADLPNKVADLEKEISNLEILG 60
      |||
DB      1 LAKKQTELEKLDNDPEKGTODELDKEAAEALDKKADLPNKVADLEKEISNLEILG 60
      |||

QY      61 GADPEDDTRALPNKATTKKAELEKTKSELDALNELGPDGDEEE 104
      |||
DB      61 GADPEDDTRALPNKATTKKAELEKTKSELDALNELGPDGDEEE 104
      |||

RESULT 2
US-08-710-749-20
; Sequence 20, Application US/08710749
; Patent No. 5955089
; GENERAL INFORMATION:
; APPLICANT: Briles, David E.
; APPLICANT: Hollingshead, Susan
; APPLICANT: Becker, Robert
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford
; STREET: 530 Fifth Avenue
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:

```

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/710,749  
FILING DATE: 20-SEP-1996  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Frommer, William S.  
REGISTRATION NUMBER: 25,506  
REFERENCE/DOCKET NUMBER: 454312-2074  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 840-3333  
TELEFAX: (212) 840-0712  
INFORMATION FOR SEQ ID NO: 20:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 104 amino acids  
TYPE: amino acid  
STRANDEDNESS: n/a  
TOPOLOGY: linear  
MOLECULE TYPE: amino acid  
US-08-710-749-20

Query Match 98.8%; Score 512; DB 2; Length 104;  
Best Local Similarity 99.0%; Pred. No. 2,6e-42;  
Matches 103; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 1 LAKKQTELEKLDNLDPBGKTQDELDEKAAAEALDKKADLPNKVADLEKEISNLEILLG 60

Db 1 LAKKQTELEKLDNLDPBGKTQDELDEKAAAEALDKKADLPNKVADLEKEISNLEILLG 60

Oy 61 GADEDDTALPNKLTATKKAELKTKQKELDPAALNELGPDGDEER 104

Db 61 GADEDDTALPNKLTATKKAELKTKQKELDPAALNELGPDGDEER 104

## RESULT 3

US-08-529-055-47

Sequence 47, Application US/08529055

Patent No. 6592876

GENERAL INFORMATION:

APPLICANT: Briles, David E.

APPLICANT: McDaniel, Larry S.

APPLICANT: Swiatlo, Edwin

APPLICANT: Yoether, Janet

APPLICANT: Brooks-Walter, Alexis

TITLE OF INVENTION: Pneumococcal Genes, Portions

TITLE OF INVENTION: Theeoff, Expression Products

TITLE OF INVENTION: Therefrom, and Uses of Such Genes,

TITLE OF INVENTION: Portions and Products

NUMBER OF SEQUENCES: 73

CORRESPONDENCE ADDRESS:

ADDRESSEE: Curtiss, Morris & Safford, P.C.

STREET: 530 Fifth Avenue

CITY: New York

STATE: NY

COUNTRY: USA

ZIP: 10036

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/529,055

FILING DATE: 15-SEP-1995

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Frommer, William S.

REGISTRATION NUMBER: 25,506

REFERENCE/DOCKET NUMBER: 454312-2400

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 840-3333  
TELEFAX: (212) 840-0712  
INFORMATION FOR SEQ ID NO: 47:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 213 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-529-055-47

Query Match 98.1%; Score 508; DB 4; Length 213;  
Best Local Similarity 98.1%; Pred. No. 1.5e-41;  
Matches 102; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy 1 LAKKQTELEKLDNLDPBGKTQDELDEKAAAEALDKKADLPNKVADLEKEISNLEILLG 60

Db 59 LAKKQTELEKLDNLDPBGKTQDELDEKAAAEALDKKADLPNKVADLEKEISNLEILLG 118

Oy 61 GADEDDTALPNKLTATKKAELKTKQKELDPAALNELGPDGDEER 104

Db 119 GADEDDTALPNKLTATKKAELKTKQKELDPAALNELGPDGDEER 162

## RESULT 4

US-08-714-741-32

Sequence 32, Application US/08714741

Patent No. 6500613

GENERAL INFORMATION:

APPLICANT: Briles, David E.

APPLICANT: McDaniel, Larry S.

APPLICANT: Swiatlo, Edwin

APPLICANT: Yoether, Janet

APPLICANT: Crain, Marilyn J.

APPLICANT: Hollingshead, Susan

APPLICANT: Tatl, Rebecca

APPLICANT: Brooks-Walter, Alexis

TITLE OF INVENTION: PNEUMOCOCCAL GENES, PORTIONS THEREOF,

TITLE OF INVENTION: EXPRESSION PRODUCTS THEREFROM, AND USES OF SUCH GENES,

TITLE OF INVENTION: PORTIONS AND PRODUCTS

NUMBER OF SEQUENCES: 47

CORRESPONDENCE ADDRESS:

ADDRESSEE: Curtiss, Morris & Safford, P.C.

STREET: 530 Fifth Avenue

CITY: New York

STATE: NY

COUNTRY: U.S.

ZIP: 10036

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/714,741

FILING DATE: 16-SEP-1996

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Frommer, William S.

REGISTRATION NUMBER: 25,506

REFERENCE/DOCKET NUMBER: 454312-2460

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 840-3333

TELEFAX: (212) 840-0712

INFORMATION FOR SEQ ID NO: 32:

SEQUENCE CHARACTERISTICS:

LENGTH: 8991 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: amino acid

US-08-714-741-32

Query Match 98.1%; Score 508; DB 4; Length 8991;  
Best Local Similarity 98.1%; Pred. No. 1.4e-39;  
Matches 102; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 LAKKQTELEKLDNDPEKGTODELDKEAAEALDKKADLPNKVADLEKEISNLEILLG 60  
DB 4625 LAKKQTELEKLDNDPEKGTODELDKEAAEALDKKADLPNKVADLEKEISNLEILLG 4684

QY 61 GADPEDDTAALPNKLTATKKALEKTKQKELDAAALNLEIGPDGDEE 104  
DB 4685 GADPEDDTAALPNKLTATKKALEKTKQKELDAAALNLEIGPDGDEE 4728

RESULT 5  
US-08-710-749-19  
; Sequence 19, Application US/08710749  
; Patent No. 5955089

; GENERAL INFORMATION:  
; APPLICANT: Brillee, David E.  
; APPLICANT: Hollingshead, Susan  
; APPLICANT: Becker, Robert  
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE  
; NUMBER OF SEQUENCES: 28  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Curtis, Morris & Safford  
; STREET: 530 Fifth Avenue  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/710,749  
FILING DATE: 20-SEP-1996  
CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:  
NAME: Frommer, William S.  
REGISTRATION NUMBER: 25,506  
REFERENCE/DOCKET NUMBER: 454312-2074  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 840-3333  
TELEFAX: (212) 840-0712

INFORMATION FOR SEQ ID NO: 19:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 104 amino acids

TYPE: amino acid  
STRANDEDNESS: n/a  
TOPOLOGY: linear

MOLECULE TYPE: amino acid  
US-08-710-749-19

Query Match 96.5%; Score 500; DB 2; Length 104;  
Best Local Similarity 97.1%; Pred. No. 3.7e-41;  
Matches 101; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 LAKKQTELEKLDNDPEKGTODELDKEAAEALDKKADLPNKVADLEKEISNLEILLG 60  
DB 1 LAKKQTELEKLDNDPEKGTODELDKEAAEALDKKADLPNKVADLEKEISNLEILLG 60

QY 61 GADPEDDTAALPNKLTATKKALEKTKQKELDAAALNLEIGPDGDEE 104  
DB 61 GADSEDDTAALPNKLTATKKALEKTKQKELDAAALNLEIGPDGDEE 104

RESULT 6  
US-09-147-875A-20  
; Sequence 20, Application US/09147875A  
; Patent No. 6638516

; GENERAL INFORMATION:  
; APPLICANT: BECKER et al.  
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS  
; FILE REFERENCE: 454312-2471  
; CURRENT APPLICATION NUMBER: US/09/147,875A  
; CURRENT FILING DATE: 1999-05-24  
; NUMBER OF SEQ ID NOS: 28  
; SOFTWARE: Patentin Ver. 2.1

SEQ ID NO 20

LENGTH: 104

TYPE: PROT

ORGANISM: Streptococcus pneumoniae  
US-09-147-875A-20

Query Match 96.5%; Score 500; DB 4; Length 104;  
Best Local Similarity 97.1%; Pred. No. 3.7e-41;  
Matches 101; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 LAKKQTELEKLDNDPEKGTODELDKEAAEALDKKADLPNKVADLEKEISNLEILLG 60  
DB 1 LAKKQTELEKLDNDPEKGTODELDKEAAEALDKKADLPNKVADLEKEISNLEILLG 60

QY 61 GADPEDDTAALPNKLTATKKALEKTKQKELDAAALNLEIGPDGDEE 104  
DB 61 GADSEDDTAALPNKLTATKKALEKTKQKELDAAALNLEIGPDGDEE 104

RESULT 7

US-08-961-083-160  
; Sequence 160, Application US/08961083  
; Patent No. 6159469

; GENERAL INFORMATION:  
; APPLICANT: Choi et. al.  
; TITLE OF INVENTION: Streptococcus pneumoniae Antigenes and Vaccines  
; NUMBER OF SEQUENCES: 452  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Human Genome Sciences, Inc.  
; STREET: 9410 Key West Avenue  
; CITY: Rockville  
; STATE: Maryland  
; COUNTRY: USA  
; ZIP: 20850

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage  
COMPUTER: HP Vectra 486/33  
OPERATING SYSTEM: MSDOS version 6.2  
SOFTWARE: ASCII Text

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/961,083  
FILING DATE:

CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Brookes, A. Anders  
REGISTRATION NUMBER: 36,373  
REFERENCE/DOCKET NUMBER: PB340P2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (301) 309-8512  
TELEFAX: (301) 309-8504

INFORMATION FOR SEQ ID NO: 160:

SEQUENCE CHARACTERISTICS:  
LENGTH: 641 amino acids

TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear

MOLECULE TYPE: protein  
US-08-961-083-160

Query Match 93.2%; Score 483; DB 3; Length 641;  
Best Local Similarity 95.2%; Pred. No. 1.5e-38;  
Matches 99; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 LAKKQTELEKLDNLDPDEGKTODELDKKAABAEIDKKADELPNKVADLEKEISNLEITLLG 60  
Db 243 LAKKQTELEKLDNLDPDEGKTODELDKKAABAEIDKKADELPNKVADLEKEISNLEITLLG 302  
QY 61 GADPEDDTALPNKLTATKKALEKTOKELDALNLELGGDDEER 104  
Db 303 GADPEDDTALPNKLTATKKALEKTOKELDALNLELGGDDEER 346

## RESULT 8

US-09-536-784-160  
; Sequence 160, Application US/09536784  
; Patent No. 6573082  
; GENERAL INFORMATION:  
; APPLICANT: Choi et. al.  
; TITLE OF INVENTION: Streptococcus pneumoniae Antigenes and Vaccines  
; NUMBER OF SEQUENCES: 452  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Human Genome Sciences, Inc.  
; STREET: 9410 Key West Avenue  
; CITY: Rockville  
; STATE: Maryland  
; COUNTRY: USA  
; ZIP: 20850  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage  
; COMPUTER: HP Vectra 486/33  
; OPERATING SYSTEM: MSDOS version 6.2  
; SOFTWARE: ASCII Text  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/536,784  
; FILING DATE: 30-Oct-1997  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/961,083  
; FILING DATE: OCT-30-1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Michelle S. Marks  
; REGISTRATION NUMBER: 41,971  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (301) 309-8504  
; TELEFAX: (301) 309-8512  
; INFORMATION FOR SEQ ID NO: 160:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 641 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; SEQUENCE DESCRIPTION: SEQ ID NO: 160:  
US-09-536-784-160

Query Match 93.2%; Score 483; DB 4; Length 641;  
Best Local Similarity 95.2%; Pred. No. 1.5e-38;  
Matches 99; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 LAKKQTELEKLDNLDPDEGKTODELDKKAABAEIDKKADELPNKVADLEKEISNLEITLLG 60  
Db 243 LAKKQTELEKLDNLDPDEGKTODELDKKAABAEIDKKADELPNKVADLEKEISNLEITLLG 302  
QY 61 GADPEDDTALPNKLTATKKALEKTOKELDALNLELGGDDEER 104  
Db 303 GADPEDDTALPNKLTATKKALEKTOKELDALNLELGGDDEER 346

## RESULT 9

US-08-529-055-44  
; Sequence 44, Application US/08529055  
; Patent No. 6592876  
; GENERAL INFORMATION:  
; APPLICANT: Briles, David E.

; APPLICANT: McDaniel, Larry S.  
; APPLICANT: Swiacki, Edwin  
; APPLICANT: Yother, Janet  
; APPLICANT: Brooks-Walter, Alexis  
; TITLE OF INVENTION: Pneumococcal Genes, Portions  
; TITLE OF INVENTION: Thereof, Expression Products  
; TITLE OF INVENTION: Therefrom, and Uses of Such Genes,  
; NUMBER OF SEQUENCES: 73  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Curtis, Morris & Safford, P.C.  
; STREET: 530 Fifth Avenue  
; CITY: New York  
; STATE: NY  
; COUNTRY: USA  
; ZIP: 10036

## COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/529,055  
; FILING DATE: 15-SEP-1995  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Frommer, William S.  
; REGISTRATION NUMBER: 25,506  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 840-0712  
; TELEFAX: (212) 840-3333  
; INFORMATION FOR SEQ ID NO: 44:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 197 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-529-055-44

Query Match 92.7%; Score 480; DB 4; Length 197;  
Best Local Similarity 93.3%; Pred. No. 6.8e-39;  
Matches 97; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 LAKKQTELEKLDNLDPDEGKTODELDKKAABAEIDKKADELPNKVADLEKEISNLEITLLG 60  
Db 22 LAKKQTELEKLDNLDPDEGKTODELDKKAABAEIDKKADELPNKVADLEKEISNLEITLLG 81  
QY 61 GADPEDDTALPNKLTATKKALEKTOKELDALNLELGGDDEER 104  
Db 82 GADPEDDTALPNKLTATKKALEKTOKELDALNLELGGDDEER 125

## RESULT 10

US-08-710-749-21  
; Sequence 21, Application US/08710749  
; Patent No. 5955089  
; GENERAL INFORMATION:  
; APPLICANT: Briles, David E.  
; APPLICANT: Hollingshead, Susan  
; APPLICANT: Becker, Robert  
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE  
; NUMBER OF SEQUENCES: 28  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Curtis, Morris & Safford  
; STREET: 530 Fifth Avenue  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/710,749  
FILING DATE: 20-SEP-1996  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Frommer, William S.  
REGISTRATION NUMBER: 25,506  
REFERENCE/DOCKET NUMBER: 454312-2074  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 840-3333  
TELEFAX: (212) 840-0712  
INFORMATION FOR SEQ ID NO: 21:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 102 amino acids  
TYPE: amino acid  
STRANDEDNESS: n/a  
TOPOLOGY: linear  
MOLECULE TYPE: amino acid  
US-08-710-749-21

Query Match 92.5%; Score 479; DB 2; Length 102;  
Best Local Similarity 97.1%; Pred. No. 3.8e-39;  
Matches 101; Conservative 0; Mismatches 1; Indels 2; Gaps 2;

Qy 1 LAKKOTLEKLDNDPEKGTODELDKEAAEAELDKKADLPNKVADLEKEISNLEILLG 60  
Db 1 LAKKOTLEKLDNDPEKGTODELDKE-AAEAELDKKADLPNKVADLEKEISNLEILLG 58  
Qy 61 GADPEDDTAALPNKLTATKKALEKTKQKELDAALNELGPDGDEE 104  
Db 59 GADSEDDTAALPNKLTATKKALEKTKQKELDAALNELGPDGDEE 102

RESULT 11  
US-09-147-875A-18  
Sequence 18, Application US/09147875A  
Patent No. 6638516  
GENERAL INFORMATION:  
APPLICANT: BECKER et al.  
TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS  
FILE REFERENCE: 454312-2471  
CURRENT APPLICATION NUMBER: US/09/147,875A  
CURRENT FILING DATE: 1999-05-24  
NUMBER OF SEQ ID NOS: 28  
SOFTWARE: Patentin Ver. 2.1  
SEQ ID NO 18  
LENGTH: 102  
TYPE: PRT  
ORGANISM: Streptococcus pneumoniae  
US-09-147-875A-18

Query Match 92.5%; Score 479; DB 4; Length 102;  
Best Local Similarity 97.1%; Pred. No. 3.8e-39;  
Matches 101; Conservative 0; Mismatches 1; Indels 2; Gaps 2;

Qy 1 LAKKOTLEKLDNDPEKGTODELDKEAAEAELDKKADLPNKVADLEKEISNLEILLG 60  
Db 1 LAKKOTLEKLDNDPEKGTODELDKE-AAEAELDKKADLPNKVADLEKEISNLEILLG 58  
Qy 61 GADPEDDTAALPNKLTATKKALEKTKQKELDAALNELGPDGDEE 104  
Db 59 GADSEDDTAALPNKLTATKKALEKTKQKELDAALNELGPDGDEE 102

RESULT 12  
US-08-529-055-52  
Sequence 52, Application US/08529055  
Patent No. 6592876  
GENERAL INFORMATION:

APPLICANT: Briles, David E.  
APPLICANT: McDaniel, Larry S.  
APPLICANT: Swiatlo, Edwin  
APPLICANT: Yocher, Janet  
APPLICANT: Brooks-Walter, Alexis  
TITLE OF INVENTION: Pneumococcal Genes, Portions  
TITLE OF INVENTION: Thereof, Expression Products  
TITLE OF INVENTION: Thereof, and Uses of Such Genes,  
Portions and Products  
NUMBER OF SEQUENCES: 73  
CORRESPONDENCE ADDRESS:  
ADDRESSER: Curtis, Morris & Safford, P.C.  
STREET: 530 Fifth Avenue  
CITY: New York  
STATE: NY  
COUNTRY: USA  
ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/529,055  
FILING DATE: 15-SEP-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Frommer, William S.  
REGISTRATION NUMBER: 25,506  
REFERENCE/DOCKET NUMBER: 454312-2400  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 840-3333  
TELEFAX: (212) 840-0712  
INFORMATION FOR SEQ ID NO: 52:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 233 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-529-055-52

Query Match 88.6%; Score 459; DB 4; Length 233;  
Best Local Similarity 90.4%; Pred. No. 8.9e-37;  
Matches 94; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

Qy 1 LAKKOTLEKLDNDPEKGTODELDKEAAEAELDKKADLPNKVADLEKEISNLEILLG 60  
Db 53 LAKKOTLEKLDNDPEKGTODELDKEAGAEELDKKADLPNKVADLEKEISNLEILLG 112  
Qy 61 GADPEDDTAALPNKLTATKKALEKTKQKELDAALNELGPDGDEE 104  
Db 113 GADSEDDTAALPNKLTATKKALEKTKQKELDAALNELGPDGDEE 156

RESULT 13  
US-08-710-749-18  
Sequence 18, Application US/08710749  
Patent No. 5955089  
GENERAL INFORMATION:  
APPLICANT: Briles, David E.  
APPLICANT: Hollingshead, Susan  
APPLICANT: Becker, Robert  
TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE  
NUMBER OF SEQUENCES: 28  
CORRESPONDENCE ADDRESS:  
ADDRESSER: Curtis, Morris & Safford  
STREET: 530 Fifth Avenue  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036

```
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/710,749
FILING DATE: 20-SEP-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Frommer, William S.
REGISTRATION NUMBER: 25,506
REFERENCE/DOCKET NUMBER: 454312-2074
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 840-3333
TELEFAX: (212) 840-0712
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 80 amino acids
TYPE: amino acid
STRANDEDNESS: n/a
TOPOLOGY: linear
MOLECULE TYPE: amino acid
US-08-710-749-18
```

```
Query Match 72.4%; Score 375; DB 2; Length 80;
Best Local Similarity 95.0%; Pred. No. 3e-29;
Matches 76; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
```

```
QY 25 LDYBAEAEIDDKKADLPNKVADLEKEISNLEIILGGADPEDDTAAIPNKLATKKAELEK 84
DB 1 LDKEAGAEAEIDDKKADGPNKYSDLEKEISNLEIILGGADSEDDTAAIPNKLATKKAELEK 60
```

```
QY 85 TOKELDAALNELGPDGDEE 104
DB 61 TOKELDAALNELGPDGDEE 80
```

## RESULT 14

```
US-09-147-875A-19
Sequence 19, Application US/09147875A
Patent No. 6638516
GENERAL INFORMATION:
APPLICANT: BECKER et al.
TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS
FILE REFERENCE: 454312-2471
CURRENT APPLICATION NUMBER: US/09/147,875A
CURRENT FILING DATE: 1999-05-24
NUMBER OF SEQ ID NOS: 28
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 19
LENGTH: 80
TYPE: PRT
ORGANISM: Streptococcus pneumoniae
US-09-147-875A-19
```

```
Query Match 72.4%; Score 375; DB 4; Length 80;
Best Local Similarity 95.0%; Pred. No. 3e-29;
Matches 76; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
```

```
QY 25 LDYBAEAEIDDKKADLPNKVADLEKEISNLEIILGGADPEDDTAAIPNKLATKKAELEK 84
DB 1 LDKEAGAEAEIDDKKADGPNKYSDLEKEISNLEIILGGADSEDDTAAIPNKLATKKAELEK 60
```

```
QY 85 TOKELDAALNELGPDGDEE 104
DB 61 TOKELDAALNELGPDGDEE 80
```

```
RESULT 15
US-08-710-749-26
Sequence 26, Application US/08710749
Patent No. 5955089
```

```
GENERAL INFORMATION:
APPLICANT: Biles, David E.
APPLICANT: Hollingshead, Susan
APPLICANT: Becker, Robert
TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE
PROTEINS
NUMBER OF SEQUENCES: 28
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Curtis, Morris & Safford
STREET: 530 Fifth Avenue
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/710,749
FILING DATE: 20-SEP-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Frommer, William S.
REGISTRATION NUMBER: 25,506
REFERENCE/DOCKET NUMBER: 454312-2074
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 840-3333
TELEFAX: (212) 840-0712
INFORMATION FOR SEQ ID NO: 26:
SEQUENCE CHARACTERISTICS:
LENGTH: 108 amino acids
TYPE: amino acid
STRANDEDNESS: n/a
TOPOLOGY: linear
MOLECULE TYPE: amino acid
US-08-710-749-26
```

```
Query Match 62.0%; Score 321; DB 2; Length 108;
Best Local Similarity 65.7%; Pred. No. 6.9e-24;
Matches 71; Conservative 11; Mismatches 22; Indels 4; Gaps 2;
```

```
QY 1 LAKKQTELEKLDNLDPKGTQDELDKKAFAEIDDKKADLPNKVADLEKEISNLEIILG 60
DB 1 LKAAAELENTLSTLPBEGKTQDELDKKAFAEALNKKVEALPNQVSELBELSKLEDNLK 60
```

```
QY 61 GADP---ED-DTAAIPNKLATKKAELEKTKQKELDAALNELGPDGDEE 104
DB 61 DAEINVEDYIKGLBEAIAITKQAELEKTPKEIDAALNELGPDGDEE 108
```

```
Search completed: November 19, 2004, 22:39:09
Job time : 19.813 secs
```

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Comugen Ltd.

OM protein - protein search, using sw model

Run on: November 19, 2004, 22:32:07 / Search time 63.3089 Seconds  
(without alignments)  
581.739 Million cell updates/sec

Title: US-10-674-755-21

Perfect score: 518  
Sequence: 1 LAKKQTELEKLDNLDPBEKG.....TQKELDALNELPGDDEE 104

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 1570615 seqs, 354127592 residues

Total number of hits satisfying chosen parameters: 1570615

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Published Applications AA:\*

```
1: /cgn2_6/ptodata/2/pubppa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/2/pubppa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubppa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubppa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubppa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubppa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubppa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/2/pubppa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubppa/US09_PUBCOMB.pep:*
10: /cgn2_6/ptodata/2/pubppa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubppa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/2/pubppa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/2/pubppa/US10_PUBCOMB.pep:*
14: /cgn2_6/ptodata/2/pubppa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/2/pubppa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/2/pubppa/US10D_PUBCOMB.pep:*
17: /cgn2_6/ptodata/2/pubppa/US10_NEW_PUB.pep:*
18: /cgn2_6/ptodata/2/pubppa/US11_NEW_PUB.pep:*
19: /cgn2_6/ptodata/2/pubppa/US60_NEW_PUB.pep:*
20: /cgn2_6/ptodata/2/pubppa/US60_PUBCOMB.pep:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	518	100.0	104	15	US-10-674-755-21
2	508	98.1	213	15	US-10-299-636-62
3	500	96.5	104	15	US-10-674-755-20
4	484	93.4	744	10	US-09-769-787-184
5	483	93.2	641	9	US-09-765-272-160
6	480	92.7	197	15	US-10-299-636-59
7	479	92.5	102	15	US-10-674-755-18
8	459	88.6	233	15	US-10-299-636-67
9	375	72.4	80	15	US-10-674-755-19
10	321	62.0	108	15	US-10-674-755-23
11	321	62.0	211	15	US-10-299-636-82
12	315	60.8	108	15	US-10-674-755-25
13	315	60.8	232	15	US-10-299-636-85

14	315	60.8	275	16	US-10-414-532-1	Sequence 1, Appl
15	315	60.8	458	15	US-10-299-636-88	Sequence 88, Appl
16	315	60.8	653	16	US-10-414-532-26	Sequence 26, Appl
17	313	60.4	108	15	US-10-674-755-24	Sequence 24, Appl
18	311	60.0	106	15	US-10-674-755-22	Sequence 22, Appl
19	306	59.1	212	15	US-10-299-636-83	Sequence 83, Appl
20	277	53.5	459	17	US-10-702-305A-18	Sequence 18, Appl
21	266	51.4	108	15	US-10-674-755-25	Sequence 25, Appl
22	259	50.0	185	15	US-10-299-636-84	Sequence 84, Appl
23	199.5	38.5	487	16	US-10-414-533-34	Sequence 34, Appl
24	199.5	38.5	524	16	US-10-414-533-21	Sequence 21, Appl
25	199.5	38.5	524	15	US-10-414-533-28	Sequence 28, Appl
26	192.5	37.2	119	15	US-10-674-755-27	Sequence 27, Appl
27	192.5	37.2	215	15	US-10-299-636-58	Sequence 58, Appl
28	187.5	36.2	290	16	US-10-414-532-65	Sequence 65, Appl
29	183	35.3	230	16	US-10-414-532-32	Sequence 32, Appl
30	183	35.3	350	16	US-10-414-533-19	Sequence 19, Appl
31	128	24.7	588	15	US-10-299-636-105	Sequence 105, Appl
32	128	24.7	588	15	US-10-299-636-96	Sequence 96, Appl
33	128	24.7	619	10	US-09-882-774-1	Sequence 1, Appl
34	128	24.7	619	15	US-10-282-122A-73702	Sequence 73702, A
35	128	24.7	619	16	US-10-414-532-72	Sequence 72, Appl
36	127	24.5	141	14	US-10-254-995-2	Sequence 2, Appl
37	127	24.5	589	9	US-09-748-875-14	Sequence 14, Appl
38	127	24.5	589	10	US-09-298-523B-14	Sequence 14, Appl
39	127	24.5	589	15	US-10-299-636-97	Sequence 97, Appl
40	127	24.5	643	15	US-10-299-636-95	Sequence 95, Appl
41	127	24.5	670	9	US-09-748-875-63	Sequence 63, Appl
42	127	24.5	670	10	US-09-298-523B-63	Sequence 63, Appl
43	127	24.5	690	9	US-09-748-875-61	Sequence 61, Appl
44	127	24.5	690	10	US-09-298-523B-61	Sequence 61, Appl
45	127	24.5	691	9	US-09-748-875-1	Sequence 1, Appl

## ALIGNMENTS

```
RESULT 1
US-10-674-755-21
; Sequence 21, Application US/10674755
; Publication No. US20040067237A1
; GENERAL INFORMATION:
; APPLICANT: BECKER et al.
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS
; FILE REFERENCE: 454312-2471
; CURRENT APPLICATION NUMBER: US/10/674,755
; PRIOR FILING DATE: 2003-09-30
; PRIOR APPLICATION NUMBER: US/09/147,875A
; PRIOR FILING DATE: 1999-05-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 21
; LENGTH: 104
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-10-674-755-21
```

```
Query Match          100.0%; Score 518; DB 15; Length 104;
Best Local Similarity 100.0%; Pred. No. 1.3e-36;
Matches 104; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 LAKKQTELEKLDNLDPBEKTODELDKEAAEELDKKADELDPNKVADLEKEISNLEILG 60
DB      1 LAKKQTELEKLDNLDPBEKTODELDKEAAEELDKKADELDPNKVADLEKEISNLEILG 60
QY      61 GADPEDTPAALPNKATTKAALEKTOKELDALNELPGDDEE 104
DB      61 GADPEDTPAALPNKATTKAALEKTOKELDALNELPGDDEE 104

RESULT 2
US-10-299-636-62
; Sequence 62, Application US/10299636
```

```
; Publication No. US20040077847A1
; GENERAL INFORMATION:
; APPLICANT: Briles, David E
; APPLICANT: McDaniel, Larry S
; APPLICANT: Swatilo, Edwin
; APPLICANT: Yocheb, Janet
; APPLICANT: Crahn, Marilyn J
; APPLICANT: Hollingshead, Susan
; APPLICANT: Tart, Rebecca
; APPLICANT: Brooke-Walter, Alexis
; TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEINS AND USES THEREOF
; FILE REFERENCE: 57909/361
; CURRENT APPLICATION NUMBER: US/10/299,636
; CURRENT FILING DATE: 2002-11-19
; PRIOR APPLICATION NUMBER: 08/714,741
; PRIOR FILING DATE: 1996-09-16
; PRIOR APPLICATION NUMBER: 08/529,055
; PRIOR FILING DATE: 1995-09-15
; NUMBER OF SEQ ID NOS: 111
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 62
; LENGTH: 213
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (2)
; OTHER INFORMATION: Xaa at position 2 is unknown
US-10-299-636-62
```

```
Query Match          98.1%; Score 508; DB 15; Length 213;
Best Local Similarity 98.1%; Pred. No. 2.1e-35;
Matches 102; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
Qy      1 LAKKQTELEKLDNDPEGKTQDELDKAAEAELDKKADDELPNKVADLEKEISNLEIILG 60
Db      59 LAKKQTELEKLDNDPEGKTQDELDKAAEAELDKKADDELPNKVADLEKEISNLEIILG 118
```

```
Qy      61 GADBEDDTAALPNKLTATKKALEKTQKELDAALNELGPDGDEE 104
Db      119 GADBEDDTAALPNKLTATKKALEKTQKELDAALNELGPDGDEE 162
```

```
RESULT 3
US-10-674-755-20
; Sequence 20, Application US/10674755
; Publication No. US20040067237A1
; GENERAL INFORMATION:
; APPLICANT: BECKER et al.
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS
; FILE REFERENCE: 454312-2471
; CURRENT APPLICATION NUMBER: US/10/674,755
; CURRENT FILING DATE: 2003-09-30
; PRIOR APPLICATION NUMBER: US/09/147,875A
; PRIOR FILING DATE: 1999-05-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 20
; LENGTH: 104
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-10-674-755-20
```

```
Query Match          96.5%; Score 500; DB 15; Length 104;
Best Local Similarity 97.1%; Pred. No. 4.3e-35;
Matches 101; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
```

```
Qy      1 LAKKQTELEKLDNDPEGKTQDELDKAAEAELDKKADDELPNKVADLEKEISNLEIILG 60
Db      1 LAKKQTELEKLDNDPEGKTQDELDKAAEAELDKKADDELPNKVADLEKEISNLEIILG 60
Qy      61 GADBEDDTAALPNKLTATKKALEKTQKELDAALNELGPDGDEE 104
```

```
Db      61 GADBEDDTAALPNKLTATKKALEKTQKELDAALNELGPDGDEE 104
RESULT 4
US-09-769-787-184
; Sequence 184, Application US/09769787
; Publication No. US20030091577A1
; GENERAL INFORMATION:
; APPLICANT: Microbial Technics Limited
; APPLICANT: Gilbert, Christophe FG
; APPLICANT: Hansbro, Philip M
; TITLE OF INVENTION: Proteins
; FILE REFERENCE: PWC/P21129WO
; CURRENT APPLICATION NUMBER: US/09/769,787
; CURRENT FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: GB 9816337.1
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: US 60/125164
; PRIOR FILING DATE: 1999-03-19
; NUMBER OF SEQ ID NOS: 388
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 184
; LENGTH: 744
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-769-787-184
```

```
Query Match          93.4%; Score 484; DB 10; Length 744;
Best Local Similarity 95.2%; Pred. No. 9.8e-33;
Matches 99; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
```

```
Qy      1 LAKKQTELEKLDNDPEGKTQDELDKAAEAELDKKADDELPNKVADLEKEISNLEIILG 60
Db      346 LAKKQTELEKLDNDPEGKTQDELDKAAEAELDKKADDELQNKVADLEKEISNLEIILG 405
```

```
Qy      61 GADBEDDTAALPNKLTATKKALEKTQKELDAALNELGPDGDEE 104
Db      406 GADBEDDTAALQNKLTATKKALEKTQKELDAALNELGPDGDEE 449
```

```
RESULT 5
US-09-765-272-160
; Sequence 160, Application US/09765272
; Patent No. US20020061545A1
; GENERAL INFORMATION:
; APPLICANT: Choi et al.
; TITLE OF INVENTION: Streptococcus pneumoniae Antigenes and Vaccines
; NUMBER OF SEQUENCES: 452
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/765,272
; FILING DATE: 22-Jan-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/961,083
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Brookes, A. Anders
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PB340P2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
```



TELEFAX: (301) 309-8512  
INFORMATION FOR SEQ ID NO: 160:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 641 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 160:  
US-09-765-272-160

Query Match 93.2%; Score 483; DB 9; Length 641;  
Best Local Similarity 95.2%; Pred. No. 1e-32;  
Matches 99; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 LAKKQTELEKLDNDLDPGKTODELDEKAAAEALDKKADLPNKVADLEKEISNLEILLG 60  
DB 243 LAKKQTELEKLDLSDPBGKTODELDKEAEALDKKADLPNKVADLEKEISNLEILLG 302  
QY 61 GADPEDDTAALPNKLTATKKALEKTOKEIDAAALNELGPDGDEE 104  
DB 303 GADSEDDTAALPNKLTATKKALEKTOKEIDAAALNELGPDGDEE 346

## RESULT 6

US-10-299-636-59  
Sequence 59, Application US/10299636  
Publication No. US20040077847A1

## GENERAL INFORMATION:

APPLICANT: Briles, David B  
APPLICANT: McDaniel, Larry S  
APPLICANT: Swiatlo, Edwin  
APPLICANT: Yother, Janet  
APPLICANT: Crain, Marilyn J  
APPLICANT: Hollingshead, Susan  
APPLICANT: Tarr, Rebecca  
APPLICANT: Brooks-Walter, Alexis  
TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEINS AND USES THEREOF  
FILE REFERENCE: 57909/361  
CURRENT APPLICATION NUMBER: US/10/299,636  
CURRENT FILING DATE: 2002-11-19  
PRIOR APPLICATION NUMBER: 08/714,741  
PRIOR FILING DATE: 1996-09-16  
PRIOR APPLICATION NUMBER: 08/529,055  
PRIOR FILING DATE: 1995-09-15  
NUMBER OF SEQ ID NOS: 111  
SOFTWARE: Patentin Ver. 2.1  
SEQ ID NO 59  
LENGTH: 197  
TYPE: prt  
ORGANISM: Streptococcus pneumoniae  
US-10-299-636-59

US-10-299-636-59

Query Match 92.7%; Score 480; DB 15; Length 197;  
Best Local Similarity 93.3%; Pred. No. 4.5e-33;  
Matches 97; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 LAKKQTELEKLDNDLDPGKTODELDEKAAAEALDKKADLPNKVADLEKEISNLEILLG 60  
DB 22 LAQKPTGLEKLDLSDPBGKTODELDKEAEGAEALDKKADLPNKVADLEKEISNLEILLG 81  
QY 61 GADPEDDTAALPNKLTATKKALEKTOKEIDAAALNELGPDGDEE 104  
DB 82 GADSEDDTAALPNKLTATKKALEKTOKEIDAAALNELGPDGDEE 125

## RESULT 7

US-10-674-755-18  
Sequence 18, Application US/10674755  
Publication No. US20040067237A1  
GENERAL INFORMATION:

APPLICANT: BECKER et al.  
TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS

FILE REFERENCE: 454312-2471  
CURRENT APPLICATION NUMBER: US/10/674,755  
CURRENT FILING DATE: 2003-09-30  
PRIOR APPLICATION NUMBER: US/09/147,875A  
PRIOR FILING DATE: 1999-05-24  
NUMBER OF SEQ ID NOS: 28  
SOFTWARE: Patentin Ver. 2.1  
SEQ ID NO 18  
LENGTH: 102  
TYPE: prt  
ORGANISM: Streptococcus pneumoniae  
US-10-674-755-18

Query Match 92.5%; Score 479; DB 15; Length 102;  
Best Local Similarity 97.1%; Pred. No. 2.5e-33;  
Matches 101; Conservative 0; Mismatches 1; Indels 2; Gaps 2;

QY 1 LAKKQTELEKLDNDLDPGKTODELDEKAAAEALDKKADLPNKVADLEKEISNLEILLG 60  
DB 1 LAKKQTELEKLDLSDPBGKTODELDKEAEGAEALDKKADLPNKVADLEKEISNLEILLG 58  
QY 61 GADPEDDTAALPNKLTATKKALEKTOKEIDAAALNELGPDGDEE 104  
DB 59 GADSEDDTAALPNKLTATKKALEKTOKEIDAAALNELGPDGDEE 102

## RESULT 8

US-10-299-636-67  
Sequence 67, Application US/10299636  
Publication No. US20040077847A1

## GENERAL INFORMATION:

APPLICANT: Briles, David B  
APPLICANT: McDaniel, Larry S  
APPLICANT: Swiatlo, Edwin  
APPLICANT: Yother, Janet  
APPLICANT: Crain, Marilyn J  
APPLICANT: Hollingshead, Susan  
APPLICANT: Tarr, Rebecca  
APPLICANT: Brooks-Walter, Alexis  
TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEINS AND USES THEREOF  
FILE REFERENCE: 57909/361  
CURRENT APPLICATION NUMBER: US/10/299,636  
CURRENT FILING DATE: 2002-11-19  
PRIOR APPLICATION NUMBER: 08/714,741  
PRIOR FILING DATE: 1996-09-16  
PRIOR APPLICATION NUMBER: 08/529,055  
PRIOR FILING DATE: 1995-09-15  
NUMBER OF SEQ ID NOS: 111  
SOFTWARE: Patentin Ver. 2.1  
SEQ ID NO 67  
LENGTH: 233  
TYPE: prt  
ORGANISM: Streptococcus pneumoniae  
US-10-299-636-67

US-10-299-636-67

Query Match 88.6%; Score 459; DB 15; Length 233;  
Best Local Similarity 90.4%; Pred. No. 3.3e-31;  
Matches 94; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

QY 1 LAKKQTELEKLDNDLDPGKTODELDEKAAAEALDKKADLPNKVADLEKEISNLEILLG 60  
DB 53 LAQKPTGLEKLDLSDPBGKTODELDKEAEGAEALDKKADLPNKVADLEKEISNLEILLG 112  
QY 61 GADPEDDTAALPNKLTATKKALEKTOKEIDAAALNELGPDGDEE 104

Db 113 GAUDEDPTAALPNKLAAXKAELEKTKELDPAAPNELGPDGDEEB 156

## RESULT 9

US-10-674-755-19  
; Sequence 19; Application US/10674755  
; Publication No. US20040067237A1  
; GENERAL INFORMATION:  
; APPLICANT: BECKER et al.  
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS  
; FILE REFERENCE: 454312-2471  
; CURRENT APPLICATION NUMBER: US/10/674,755  
; CURRENT FILING DATE: 2003-09-30  
; PRIOR APPLICATION NUMBER: US/09/147,875A  
; PRIOR FILING DATE: 1999-05-24  
; NUMBER OF SEQ ID NOS: 28  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 19  
; LENGTH: 80  
; TYPE: PRT  
; ORGANISM: Streptococcus pneumoniae  
US-10-674-755-19

Query Match 72.4%; Score 375; DB 15; Length 80;  
Best Local Similarity 95.0%; Pred. No. 1.3e-24;  
Matches 76; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 25 LDKKAFAEELDKRADELPNKVADLEKEISNLEILLGADDEDDPAALPNKLTATKKALEK 84

Db 1 LDKKAGAEELDKKADGLPNKVSDEKEISNLEILLGADSEDDPTAALPNKLTATKKALEK 60

OY 85 TOKELDAALNELGPDGDEEB 104

Db 61 TOKELDAALNELGPDGDEEB 80

## RESULT 10

US-10-674-755-23  
; Sequence 23; Application US/10674755  
; Publication No. US20040067237A1  
; GENERAL INFORMATION:  
; APPLICANT: BECKER et al.  
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS  
; FILE REFERENCE: 454312-2471  
; CURRENT APPLICATION NUMBER: US/10/674,755  
; CURRENT FILING DATE: 2003-09-30  
; PRIOR APPLICATION NUMBER: US/09/147,875A  
; PRIOR FILING DATE: 1999-05-24  
; NUMBER OF SEQ ID NOS: 28  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 23  
; LENGTH: 108  
; TYPE: PRT  
; ORGANISM: Streptococcus pneumoniae  
US-10-674-755-23

Query Match 62.0%; Score 321; DB 15; Length 108;  
Best Local Similarity 65.7%; Pred. No. 6.9e-20;  
Matches 71; Conservative 11; Mismatches 22; Indels 4; Gaps 2;

OY 1 LAKKQTELEKLDNLDPBEGKTODELDKKAFAEELDKKADDELPNKVADLEKEISNLEILLG 60

Db 1 LKKAFAELLENLSTLDPBEGKTODELDKKAFAEELINKVVALPNQVSELEBELSKLEEDNLK 60

OY 61 GADP---ED-DTAAALPNKLTATKKALEKTKELDPAALNELGPDGDEEB 104

Db 61 DAETNNVEDYIKGELBEAIAITKQALEKTPKELDPAALNELGPDGDEEB 108

## RESULT 11

US-10-299-636-82  
; Sequence 82; Application US/10299636  
; Publication No. US2004007847A1

; GENERAL INFORMATION:

; APPLICANT: Brillee, David E  
; APPLICANT: McDaniel, Larry S  
; APPLICANT: Swiatlo, Edwin  
; APPLICANT: Yocher, Janet  
; APPLICANT: Crain, Marilyn J  
; APPLICANT: Hollingshead, Susan  
; APPLICANT: Tarr, Rebecca  
; APPLICANT: Brooks-Walter, Alexis  
; TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEINS AND USES THEREOF  
; FILE REFERENCE: 57909/361  
; CURRENT APPLICATION NUMBER: US/10/299,636  
; CURRENT FILING DATE: 2002-11-19  
; PRIOR APPLICATION NUMBER: 08/714,741  
; PRIOR FILING DATE: 1996-09-16  
; PRIOR APPLICATION NUMBER: 08/529,055  
; PRIOR FILING DATE: 1995-09-15  
; NUMBER OF SEQ ID NOS: 111  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 82  
; LENGTH: 211  
; TYPE: PRT  
; ORGANISM: Streptococcus pneumoniae  
US-10-299-636-82

Query Match 62.0%; Score 321; DB 15; Length 211;  
Best Local Similarity 65.7%; Pred. No. 1.5e-19;  
Matches 71; Conservative 11; Mismatches 22; Indels 4; Gaps 2;

OY 1 LAKKQTELEKLDNLDPBEGKTODELDKKAFAEELDKKADDELPNKVADLEKEISNLEILLG 60

Db 25 LKKAFAELLENLSTLDPBEGKTODELDKKAFAEELINKVVALPNQVSELEBELSKLEEDNLK 84

OY 61 GADP---ED-DTAAALPNKLTATKKALEKTKELDPAALNELGPDGDEEB 104

Db 85 DAETNNVEDYIKGELBEAIAITKQALEKTPKELDPAALNELGPDGDEEB 132

## RESULT 12

US-10-674-755-25  
; Sequence 25; Application US/10674755  
; Publication No. US20040067237A1  
; GENERAL INFORMATION:  
; APPLICANT: BECKER et al.  
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS  
; FILE REFERENCE: 454312-2471  
; CURRENT APPLICATION NUMBER: US/10/674,755  
; CURRENT FILING DATE: 2003-09-30  
; PRIOR APPLICATION NUMBER: US/09/147,875A  
; PRIOR FILING DATE: 1999-05-24  
; NUMBER OF SEQ ID NOS: 28  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 25  
; LENGTH: 108  
; TYPE: PRT  
; ORGANISM: Streptococcus pneumoniae  
US-10-674-755-25

Query Match 60.8%; Score 315; DB 15; Length 108;  
Best Local Similarity 65.7%; Pred. No. 2.2e-19;  
Matches 71; Conservative 11; Mismatches 22; Indels 4; Gaps 2;

OY 1 LAKKQTELEKLDNLDPBEGKTODELDKKAFAEELDKKADDELPNKVADLEKEISNLEILLG 60

Db 1 LKKAFAELLENLSTLDPBEGKTODELDKKAFAEELINKVVALPNQVSELEBELSKLEEDNLK 60

OY 61 GADP---ED-DTAAALPNKLTATKKALEKTKELDPAALNELGPDGDEEB 104

Db 61 DAETNNVEDYIKGELBEAIAITKQALEKTPKELDPAALNELGPDGDEEB 108

## RESULT 13

US-10-299-636-85

; Sequence 85, Application US/10299636  
; Publication No. US20040077847A1  
; GENERAL INFORMATION:  
; APPLICANT: Briles, David E  
; APPLICANT: McDaniel, Larry S  
; APPLICANT: Swiatlo, Edwin  
; APPLICANT: Yotter, Janet  
; APPLICANT: Crain, Marilyn J  
; APPLICANT: Hollingshead, Susan  
; APPLICANT: Tarr, Rebecca  
; APPLICANT: Brooks-Walter, Alexis  
; TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEINS AND USES THEREOF  
; FILE REFERENCE: 57909/361  
; CURRENT FILING DATE: 2002-11-19  
; PRIOR APPLICATION NUMBER: US/10/299,636  
; PRIOR FILING DATE: 1996-09-16  
; PRIOR APPLICATION NUMBER: 08/529,055  
; PRIOR FILING DATE: 1995-09-15  
; NUMBER OF SEQ ID NOS: 111  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 85  
; LENGTH: 232  
; TYPE: PRT  
; ORGANISM: Streptococcus pneumoniae  
US-10-299-636-85

Query Match 60.8%; Score 315; DB 15; Length 232;  
Best Local Similarity 65.7%; Pred. No. 5.5e-19;  
Matches 71; Conservative 11; Mismatches 22; Indels 4; Gaps 2;

Qy 1 LAKKQTELEKLDNDPPEKGTODELDKEAAEALDKKADELPNKVADLEKEISNLEILLG 60  
Db 51 LEDALELEKVLATDPEKGTODELDKEAAEALNEKVEALONQVALEEBELSKLEDNLIK 110  
Qy 61 GADP---ED-DTAAIPNKLATKKALEKTKQKELDAALNELGPDGDEEE 104  
Db 111 DAEITNNVEDYIKKGLBAIATKKALEKTKQKELDAALNELGPDGDEEE 158

RESULT 14  
US-10-414-532-1  
; Sequence 1, Application US/10414532  
; Publication No. US20040101531A1  
; GENERAL INFORMATION:  
; APPLICANT: KANG, HO YOUNG  
; APPLICANT: CURTIS III, ROY  
; TITLE OF INVENTION: IMPROVED IMMUNOGENIC COMPOSITIONS AND VACCINES COMPRISING  
; TITLE OF INVENTION: CARRIER BACTERIA THAT SECRETE ANTIGENS  
; FILE REFERENCE: 56029-40437  
; CURRENT FILING DATE: 2003-04-15  
; PRIOR APPLICATION NUMBER: 60/372,710  
; PRIOR FILING DATE: 2002-04-16  
; NUMBER OF SEQ ID NOS: 72  
; SOFTWARE: PatentIn Ver. 3.2  
; SEQ ID NO 1  
; LENGTH: 275  
; TYPE: PRT  
; ORGANISM: Streptococcus pneumoniae  
US-10-414-532-1

Query Match 60.8%; Score 315; DB 16; Length 275;  
Best Local Similarity 65.7%; Pred. No. 6.7e-19;  
Matches 71; Conservative 11; Mismatches 22; Indels 4; Gaps 2;

Qy 1 LAKKQTELEKLDNDPPEKGTODELDKEAAEALDKKADELPNKVADLEKEISNLEILLG 60  
Db 167 LEDALELEKVLATDPEKGTODELDKEAAEALNEKVEALONQVALEEBELSKLEDNLIK 226  
Qy 61 GADP---ED-DTAAIPNKLATKKALEKTKQKELDAALNELGPDGDEEE 104  
Db 227 DAEITNNVEDYIKKGLBAIATKKALEKTKQKELDAALNELGPDGDEEE 274

RESULT 15  
US-10-299-636-88  
; Sequence 88, Application US/10299636  
; Publication No. US20040077847A1  
; GENERAL INFORMATION:  
; APPLICANT: Briles, David E  
; APPLICANT: McDaniel, Larry S  
; APPLICANT: Swiatlo, Edwin  
; APPLICANT: Yotter, Janet  
; APPLICANT: Crain, Marilyn J  
; APPLICANT: Hollingshead, Susan  
; APPLICANT: Tarr, Rebecca  
; APPLICANT: Brooks-Walter, Alexis  
; TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEINS AND USES THEREOF  
; FILE REFERENCE: 57909/361  
; CURRENT FILING DATE: 2002-11-19  
; PRIOR APPLICATION NUMBER: US/10/299,636  
; PRIOR FILING DATE: 1996-09-16  
; PRIOR APPLICATION NUMBER: 08/529,055  
; PRIOR FILING DATE: 1995-09-15  
; NUMBER OF SEQ ID NOS: 111  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 88  
; LENGTH: 458  
; TYPE: PRT  
; ORGANISM: Streptococcus pneumoniae  
US-10-299-636-88

Query Match 60.8%; Score 315; DB 15; Length 458;  
Best Local Similarity 65.7%; Pred. No. 1.2e-18;  
Matches 71; Conservative 11; Mismatches 22; Indels 4; Gaps 2;

Qy 1 LAKKQTELEKLDNDPPEKGTODELDKEAAEALDKKADELPNKVADLEKEISNLEILLG 60  
Db 276 LEDALELEKVLATDPEKGTODELDKEAAEALNEKVEALONQVALEEBELSKLEDNLIK 335  
Qy 61 GADP---ED-DTAAIPNKLATKKALEKTKQKELDAALNELGPDGDEEE 104  
Db 336 DAEITNNVEDYIKKGLBAIATKKALEKTKQKELDAALNELGPDGDEEE 383

Search completed: November 19, 2004, 23:28:07  
Job time : 64.3089 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 19, 2004, 22:08:34 ; Search time 240.024 Seconds  
(without alignments)  
480.084 Million cell updates/sec

Title: US-10-674-755-21

Perfect score: 518  
Sequence: 1 LAKKQTELEKLDNLDPECK.....TQKELDAALNEKGPDDERE 104

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 6730630 seqs, 1107998698 residues

Total number of hits satisfying chosen parameters: 6730630

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : Pending Patents AA Main:\*

1: /cgn2\_6/ptodata/1/paa/PCTUS\_COMB.pep.\*  
2: /cgn2\_6/ptodata/1/paa/US06\_COMB.pep.\*  
3: /cgn2\_6/ptodata/1/paa/US07\_COMB.pep.\*  
4: /cgn2\_6/ptodata/1/paa/US08\_COMB.pep.\*  
5: /cgn2\_6/ptodata/1/paa/US081\_COMB.pep.\*  
6: /cgn2\_6/ptodata/1/paa/US082\_COMB.pep.\*  
7: /cgn2\_6/ptodata/1/paa/US083\_COMB.pep.\*  
8: /cgn2\_6/ptodata/1/paa/US084\_COMB.pep.\*  
9: /cgn2\_6/ptodata/1/paa/US085\_COMB.pep.\*  
10: /cgn2\_6/ptodata/1/paa/US086\_COMB.pep.\*  
11: /cgn2\_6/ptodata/1/paa/US087\_COMB.pep.\*  
12: /cgn2\_6/ptodata/1/paa/US088\_COMB.pep.\*  
13: /cgn2\_6/ptodata/1/paa/US089\_COMB.pep.\*  
14: /cgn2\_6/ptodata/1/paa/US090\_COMB.pep.\*  
15: /cgn2\_6/ptodata/1/paa/US091\_COMB.pep.\*  
16: /cgn2\_6/ptodata/1/paa/US092\_COMB.pep.\*  
17: /cgn2\_6/ptodata/1/paa/US093\_COMB.pep.\*  
18: /cgn2\_6/ptodata/1/paa/US094\_COMB.pep.\*  
19: /cgn2\_6/ptodata/1/paa/US095\_COMB.pep.\*  
20: /cgn2\_6/ptodata/1/paa/US096\_COMB.pep.\*  
21: /cgn2\_6/ptodata/1/paa/US097A\_COMB.pep.\*  
22: /cgn2\_6/ptodata/1/paa/US097B\_COMB.pep.\*  
23: /cgn2\_6/ptodata/1/paa/US098\_COMB.pep.\*  
24: /cgn2\_6/ptodata/1/paa/US099A\_COMB.pep.\*  
25: /cgn2\_6/ptodata/1/paa/US099B\_COMB.pep.\*  
26: /cgn2\_6/ptodata/1/paa/US100\_COMB.pep.\*  
27: /cgn2\_6/ptodata/1/paa/US101\_COMB.pep.\*  
28: /cgn2\_6/ptodata/1/paa/US102\_COMB.pep.\*  
29: /cgn2\_6/ptodata/1/paa/US103\_COMB.pep.\*  
30: /cgn2\_6/ptodata/1/paa/US104\_COMB.pep.\*  
31: /cgn2\_6/ptodata/1/paa/US105\_COMB.pep.\*  
32: /cgn2\_6/ptodata/1/paa/US106\_COMB.pep.\*  
33: /cgn2\_6/ptodata/1/paa/US107\_COMB.pep.\*  
34: /cgn2\_6/ptodata/1/paa/US108\_COMB.pep.\*  
35: /cgn2\_6/ptodata/1/paa/US109\_COMB.pep.\*  
36: /cgn2\_6/ptodata/1/paa/US60\_COMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	518	100.0	104	32	US-10-674-755-21	Sequence 21, Appl
2	508	98.1	213	28	US-10-299-636-62	Sequence 62, Appl
3	500	96.5	104	32	US-10-674-755-20	Sequence 20, Appl
4	498	96.1	502	22	US-09-791-537-114569	Sequence 114569,
5	484	93.4	744	1	PCT-US03-27401-251	Sequence 251, App
6	484	93.4	744	22	US-09-765-787-184	Sequence 184, App
7	484	93.4	744	30	US-10-472-928-32	Sequence 32, Appl
8	484	93.4	744	34	US-10-873-528A-184	Sequence 184, App
9	484	93.4	744	34	US-10-873-528A-184	Sequence 184, App
10	483	93.2	641	13	US-08-961-083-160	Sequence 160, App
11	483	93.2	641	22	US-09-765-271-160	Sequence 160, App
12	483	93.2	641	22	US-09-765-272-160	Sequence 160, App
13	483	93.2	641	22	US-09-765-272A-160	Sequence 160, App
14	480	92.7	107	32	US-10-299-636-59	Sequence 59, Appl
15	479	92.5	102	32	US-10-674-755-18	Sequence 18, Appl
16	459	88.6	233	28	US-10-299-636-67	Sequence 67, Appl
17	375	72.4	80	32	US-10-674-755-19	Sequence 19, Appl
18	322	62.2	479	22	US-09-791-537-114578	Sequence 114578,
19	321	62.0	108	32	US-10-674-755-23	Sequence 23, Appl
20	321	62.0	211	28	US-10-299-636-82	Sequence 82, Appl
21	315	60.8	108	32	US-10-674-755-25	Sequence 25, Appl
22	315	60.8	232	28	US-10-299-636-85	Sequence 85, Appl
23	315	60.8	275	30	US-10-414-532-1	Sequence 1, Appl
24	315	60.8	275	36	US-60-372-710-1	Sequence 1, Appl
25	315	60.8	458	28	US-10-299-636-88	Sequence 88, Appl
26	315	60.8	653	30	US-10-414-532-26	Sequence 26, Appl
27	315	60.8	653	36	US-60-372-710-26	Sequence 26, Appl
28	313	60.4	108	32	US-10-674-755-24	Sequence 24, Appl
29	311	60.0	106	32	US-10-674-755-22	Sequence 22, Appl
30	306	59.1	212	28	US-10-299-636-83	Sequence 83, Appl
31	297	57.3	480	22	US-09-791-537-114577	Sequence 114577,
32	277	53.5	459	1	PCT-US03-35529-18	Sequence 18, Appl
33	277	53.5	459	33	US-10-702-305A-18	Sequence 18, Appl
34	266	51.4	108	32	US-10-674-755-26	Sequence 26, Appl
35	259	50.0	185	28	US-10-299-636-84	Sequence 84, Appl
36	245.5	47.4	256	22	US-09-791-537-136618	Sequence 136618,
37	245.5	47.4	461	22	US-09-791-537-114571	Sequence 114571,
38	199.5	38.5	487	1	PCT-US03-11802-21	Sequence 21, Appl
39	199.5	38.5	487	30	US-10-414-532-34	Sequence 34, Appl
40	199.5	38.5	487	30	US-10-414-532-31	Sequence 31, Appl
41	199.5	38.5	524	30	US-10-414-532-28	Sequence 28, Appl
42	199.5	38.5	524	36	US-60-372-710-28	Sequence 28, Appl
43	192.5	37.2	119	32	US-10-674-755-27	Sequence 27, Appl
44	192.5	37.2	215	28	US-10-299-636-58	Sequence 58, Appl
45	187.5	36.2	290	30	US-10-414-532-65	Sequence 65, Appl

## ALIGNMENTS

RESULT 1  
US-10-674-755-21

Sequence 21, Application US/10674755

GENERAL INFORMATION:

APPLICANT: BECKER et al.

TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS

FILE REFERENCE: 454312-2471

CURRENT FILING DATE: 2003-09-30

PRIOR APPLICATION NUMBER: US/09/147,875A

NUMBER OF SEQ ID NOS: 28

SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 21

LENGTH: 104

TYPE: PRT

ORGANISM: Streptococcus pneumoniae

US-10-674-755-21

Query Match 100.0% Score 518; DB 32; Length 104;

Best Local Similarity 100.0%; Pred. No. 7.5e-38;  
Matches 104; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LAKKQTELEKLDNLDPGKQDDELKKAABAEIDKKADELPNKVADLEKEISNLEITLLG 60  
Db 1 LAKKQTELEKLDNLDPGKQDDELKKAABAEIDKKADELPNKVADLEKEISNLEITLLG 60

QY 61 GADEDDTALPNKLTATKKALEKTOKELDALNLELPGDGEER 104  
Db 61 GADEDDTALPNKLTATKKALEKTOKELDALNLELPGDGEER 104

RESULT 2  
US-10-299-636-62  
Sequence 62, Application US/10299636

GENERAL INFORMATION:  
APPLICANT: Briles, David E  
APPLICANT: McDaniel, Larry S  
APPLICANT: Swatilo, Edwin  
APPLICANT: Yochev, Janet  
APPLICANT: Crahn, Marilyn J  
APPLICANT: Hollingshead, Susan  
APPLICANT: Tart, Rebecca  
APPLICANT: Brooks-Walter, Alexis  
TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEINS AND USES THEREOF  
FILE REFERENCE: 57909/361  
CURRENT APPLICATION NUMBER: US/10/299,636  
CURRENT FILING DATE: 2002-11-19  
PRIOR APPLICATION NUMBER: 08/714,741  
PRIOR FILING DATE: 1996-09-16  
PRIOR APPLICATION NUMBER: 08/529,055  
PRIOR FILING DATE: 1995-09-15  
NUMBER OF SEQ ID NOS: 111  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 62  
LENGTH: 213  
TYPE: PRT  
ORGANISM: Streptococcus pneumoniae  
FEATURE:  
NAME/KEY: UNSURE  
LOCATION: (2)  
OTHER INFORMATION: Xaa at position 2 is unknown  
US-10-299-636-62

Query Match 98.1%; Score 508; DB 28; Length 213;  
Best Local Similarity 98.1%; Pred. No. 1.5e-36;  
Matches 102; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 LAKKQTELEKLDNLDPGKQDDELKKAABAEIDKKADELPNKVADLEKEISNLEITLLG 60  
Db 59 LAKKQTELEKLDNLDPGKQDDELKKAABAEIDKKADELPNKVADLEKEISNLEITLLG 118

QY 61 GADEDDTALPNKLTATKKALEKTOKELDALNLELPGDGEER 104  
Db 119 GADEDDTALPNKLTATKKALEKTOKELDALNLELPGDGEER 162

RESULT 3  
US-10-674-755-20  
Sequence 20, Application US/10674755  
GENERAL INFORMATION:  
APPLICANT: BECKER et al.  
TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS  
FILE REFERENCE: 454312-2471  
CURRENT APPLICATION NUMBER: US/10/674,755  
CURRENT FILING DATE: 2003-09-30  
PRIOR APPLICATION NUMBER: US/09/147,875A  
NUMBER OF SEQ ID NOS: 28  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 20  
LENGTH: 104  
TYPE: PRT

ORGANISM: Streptococcus pneumoniae  
US-10-674-755-20

Query Match 96.5%; Score 500; DB 32; Length 104;  
Best Local Similarity 97.1%; Pred. No. 2.9e-36;  
Matches 101; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 LAKKQTELEKLDNLDPGKQDDELKKAABAEIDKKADELPNKVADLEKEISNLEITLLG 60  
Db 1 LAKKQTELEKLDNLDPGKQDDELKKAABAEIDKKADELPNKVADLEKEISNLEITLLG 60

QY 61 GADEDDTALPNKLTATKKALEKTOKELDALNLELPGDGEER 104  
Db 61 GADEDDTALPNKLTATKKALEKTOKELDALNLELPGDGEER 104

RESULT 4  
US-09-791-537-114569  
Sequence 114569, Application US/09791537

GENERAL INFORMATION:  
APPLICANT: Biomomix, Inc.  
APPLICANT: Debe, Derek  
APPLICANT: Danzer, Joseph  
TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS  
METHODS OF USE THEREOF  
FILE REFERENCE: 261/210  
CURRENT APPLICATION NUMBER: US/09/791,537  
CURRENT FILING DATE: 2001-02-22  
NUMBER OF SEQ ID NOS: 153055  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 114569  
LENGTH: 502  
TYPE: PRT  
ORGANISM: Streptococcus pneumoniae  
US-09-791-537-114569

Query Match 96.1%; Score 498; DB 22; Length 502;  
Best Local Similarity 97.1%; Pred. No. 3.8e-35;  
Matches 101; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 LAKKQTELEKLDNLDPGKQDDELKKAABAEIDKKADELPNKVADLEKEISNLEITLLG 60  
Db 347 LAKKQTELEKLDNLDPGKQDDELKKAABAEIDKKADELPNKVADLEKEISNLEITLLG 406

QY 61 GADEDDTALPNKLTATKKALEKTOKELDALNLELPGDGEER 104  
Db 407 GADEDDTALPNKLTATKKALEKTOKELDALNLELPGDGEER 450

RESULT 5  
PCT-US03-27401-251  
Sequence 251, Application PC/TUS0327401  
GENERAL INFORMATION:  
APPLICANT: TUFTS UNIVERSITY  
TITLE OF INVENTION: STREPTOCOCCUS PNEUMONIAE FOR DIAGNOSIS, TREATMENT AND PREVENTION OF ACTIVE INFECTION  
FILE REFERENCE: 700355-52941-PCT  
CURRENT APPLICATION NUMBER: PCT/US03/27401  
CURRENT FILING DATE: 2003-09-02  
PRIOR APPLICATION NUMBER: US 60/407,082  
PRIOR FILING DATE: 2002-08-30  
NUMBER OF SEQ ID NOS: 560  
SOFTWARE: PatentIn version 3.2  
SEQ ID NO 251  
LENGTH: 744  
TYPE: PRT  
ORGANISM: Streptococcus pneumoniae  
PCT-US03-27401-251

Query Match 93.4%; Score 484; DB 1; Length 744;  
Best Local Similarity 95.2%; Pred. No. 1.1e-33;  
Matches 99; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

```
Query Match      93.4%; Score 484; DB 30; Length 744;
Best Local Similarity 95.2%; Pred. No. 1.1e-33;
Matches 99; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
```

Query Match	93.4%;	Score 484;	DB 34;	Length 744;
Best Local Similarity	95.2%;	Pred. No. 1.1e-33;		

Matches 99; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 LAKKQTELEKLDLNDPEGKTQDELDKBAEAELDKKADLPNKVADLEKISNLEIILG 60  
DB 346 LAKKQTELEKLDLNDPEGKTQDELDKBAEAELDKKADLPNKVADLEKISNLEIILG 405

QY 61 GADPEDDTAALPNKLTATKKALEKTKQKELDPAALNELGPDGDEE 104  
DB 406 GADPEDDTAALPNKLTATKKALEKTKQKELDPAALNELGPDGDEE 449

RESULT 10  
US-08-961-083-160  
Sequence 160, Application US/08961083  
GENERAL INFORMATION:  
APPLICANT: Choi et. al.  
TITLE OF INVENTION: Streptococcus pneumoniae Antigenes and Vaccines  
NUMBER OF SEQUENCES: 452  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Human Genome Sciences, Inc.  
STREET: 9410 Key West Avenue  
CITY: Rockville  
STATE: Maryland  
COUNTRY: USA  
ZIP: 20850  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage  
COMPUTER: HP Vectra 486/33  
OPERATING SYSTEM: MSDOS version 6.2  
SOFTWARE: ASCII Text  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/961,083  
FILING DATE: 30-Oct-1997  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: <Unknown>  
FILING DATE: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Brookes, A. Anders  
REGISTRATION NUMBER: 36,373  
REFERENCE/DOCKET NUMBER: PB340P2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (301) 309-8504  
TELEFAX: (301) 309-8512  
INFORMATION FOR SEQ ID NO: 160:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 641 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 160:  
US-08-961-083-160

Query Match 93.2%; Score 483; DB 13; Length 641;  
Best Local Similarity 95.2%; Pred. No. 1,1e-33;  
Matches 99; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 LAKKQTELEKLDLNDPEGKTQDELDKBAEAELDKKADLPNKVADLEKISNLEIILG 60  
DB 243 LAKKQTELEKLDLNDPEGKTQDELDKBAEAELDKKADLPNKVADLEKISNLEIILG 302

QY 61 GADPEDDTAALPNKLTATKKALEKTKQKELDPAALNELGPDGDEE 104  
DB 303 GADPEDDTAALPNKLTATKKALEKTKQKELDPAALNELGPDGDEE 346

RESULT 11  
US-09-765-271-160  
Sequence 160, Application US/09765271  
GENERAL INFORMATION:  
APPLICANT: Choi et. al.  
TITLE OF INVENTION: Streptococcus pneumoniae Antigenes and Vaccines

NUMBER OF SEQUENCES: 452  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Human Genome Sciences, Inc.  
STREET: 9410 Key West Avenue  
CITY: Rockville  
STATE: Maryland  
COUNTRY: USA  
ZIP: 20850  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage  
COMPUTER: HP Vectra 486/33  
OPERATING SYSTEM: MSDOS version 6.2  
SOFTWARE: ASCII Text  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/765,271  
FILING DATE: 22-Jan-2001  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 09/536,784  
FILING DATE: <Unknown>  
APPLICATION NUMBER: 08/961,083  
FILING DATE: OCT-30-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Michelle S. Marks  
REGISTRATION NUMBER: 41,971  
REFERENCE/DOCKET NUMBER: PB340P3  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (301) 309-8504  
TELEFAX: (301) 309-8512  
INFORMATION FOR SEQ ID NO: 160:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 641 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 160:  
US-09-765-271-160

Query Match 93.2%; Score 483; DB 22; Length 641;  
Best Local Similarity 95.2%; Pred. No. 1,1e-33;  
Matches 99; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 LAKKQTELEKLDLNDPEGKTQDELDKBAEAELDKKADLPNKVADLEKISNLEIILG 60  
DB 243 LAKKQTELEKLDLNDPEGKTQDELDKBAEAELDKKADLPNKVADLEKISNLEIILG 302

QY 61 GADPEDDTAALPNKLTATKKALEKTKQKELDPAALNELGPDGDEE 104  
DB 303 GADPEDDTAALPNKLTATKKALEKTKQKELDPAALNELGPDGDEE 346

RESULT 12  
US-09-765-272-160  
Sequence 160, Application US/09765272  
GENERAL INFORMATION:  
APPLICANT: Choi et. al.  
TITLE OF INVENTION: Streptococcus pneumoniae Antigenes and Vaccines  
NUMBER OF SEQUENCES: 452  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Human Genome Sciences, Inc.  
STREET: 9410 Key West Avenue  
CITY: Rockville  
STATE: Maryland  
COUNTRY: USA  
ZIP: 20850  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage  
COMPUTER: HP Vectra 486/33  
OPERATING SYSTEM: MSDOS version 6.2  
SOFTWARE: ASCII Text  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/765,272



FILING DATE: 22-Jan-2001  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/961,083  
FILING DATE: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Brookes, A. Anders  
REGISTRATION NUMBER: 36,373  
REFERENCE/DOCKET NUMBER: PB340P2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (301) 309-8504  
TELEFAX: (301) 309-8512  
INFORMATION FOR SEQ ID NO: 160:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 641 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 160:  
US-09-765-272-160

Query Match 93.2%; Score 483; DB 22; Length 641;  
Best Local Similarity 95.2%; Pred. No. 1.1e-33;  
Matches 99; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 LAKQTELEKLDNDLDPGKTQDELDEKAEAEALDKKADLPNKVADLEKEISNLEILLG 60  
DB 243 LAKQTELEKLDLSDLPBGKTQDELDEKAEAEALDKKADLPNKVADLEKEISNLEILLG 302

QY 61 GADPEDDTAALPNKLTATKKALEKTKQKELDAALNLELGGDDEE 104  
DB 303 GADXBDDTAALQNKLTATKKALEKTKQKELDAALNLELGGDDEE 346

RESULT 13  
US-09-765-272A-160  
Sequence 160, Application US/09765272A  
GENERAL INFORMATION:  
APPLICANT: Choi et. al.  
TITLE OF INVENTION: Streptococcus pneumoniae Antigens and  
Vaccines  
NUMBER OF SEQUENCES: 454  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Human Genome Sciences, Inc.  
STREET: 9410 Key West Avenue  
CITY: Rockville  
STATE: Maryland  
COUNTRY: USA  
ZIP: 20850  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 1.4MB storage  
COMPUTER: Dell Latitude C610  
OPERATING SYSTEM: Windows 2000  
SOFTWARE: ASCII Text  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/765,272A  
FILING DATE: 22-Jan-2001  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/961,083  
FILING DATE: OCT-30-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Lin J. Hymel  
REGISTRATION NUMBER: 45,414  
REFERENCE/DOCKET NUMBER: PB340P2C2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (301) 610-5790  
TELEFAX: (301) 309-8439  
INFORMATION FOR SEQ ID NO: 160:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 641 amino acids  
TYPE: amino acid

STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 160:  
US-09-765-272A-160

Query Match 93.2%; Score 483; DB 22; Length 641;  
Best Local Similarity 95.2%; Pred. No. 1.1e-33;  
Matches 99; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 LAKQTELEKLDNDLDPGKTQDELDEKAEAEALDKKADLPNKVADLEKEISNLEILLG 60  
DB 243 LAKQTELEKLDLSDLPBGKTQDELDEKAEAEALDKKADLPNKVADLEKEISNLEILLG 302

QY 61 GADPEDDTAALPNKLTATKKALEKTKQKELDAALNLELGGDDEE 104  
DB 303 GADXBDDTAALQNKLTATKKALEKTKQKELDAALNLELGGDDEE 346

RESULT 14  
US-10-299-636-59  
Sequence 59, Application US/10299636  
GENERAL INFORMATION:  
APPLICANT: Briles, David E  
APPLICANT: McDaniel, Larry S  
APPLICANT: Swiatlo, Edwin  
APPLICANT: Yotner, Janet  
APPLICANT: Crain, Marilyn J  
APPLICANT: Hollingshead, Susan  
APPLICANT: Tart, Rebecca  
APPLICANT: Brooks-Walter, Alexis  
TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEINS AND USES THEREOF  
FILE REFERENCE: 57909/361  
CURRENT APPLICATION NUMBER: US/10/299,636  
CURRENT FILING DATE: 2002-11-19  
PRIOR APPLICATION NUMBER: 08/714,741  
PRIOR FILING DATE: 1996-09-16  
PRIOR APPLICATION NUMBER: 08/529,055  
PRIOR FILING DATE: 1995-09-15  
NUMBER OF SEQ ID NOS: 111  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 59  
LENGTH: 197  
TYPE: PRT  
ORGANISM: Streptococcus pneumoniae  
US-10-299-636-59

Query Match 92.7%; Score 480; DB 28; Length 197;  
Best Local Similarity 93.3%; Pred. No. 4.2e-34;  
Matches 97; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 LAKQTELEKLDNDLDPGKTQDELDEKAEAEALDKKADLPNKVADLEKEISNLEILLG 60  
DB 22 LAKQTELEKLDLSDLPBGKTQDELDEKAEAEALDKKADLPNKVADLEKEISNLEILLG 81

QY 61 GADPEDDTAALPNKLTATKKALEKTKQKELDAALNLELGGDDEE 104  
DB 82 GADXBDDTAALPNKLTATKKALEKTKQKELDAALNLELGGDDEE 125

RESULT 15  
US-10-674-755-18  
Sequence 18, Application US/10674755  
GENERAL INFORMATION:  
APPLICANT: BECKER et al.  
TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE PROTEINS  
FILE REFERENCE: 454312-2471  
CURRENT FILING DATE: 2003-09-30  
CURRENT APPLICATION NUMBER: US/10/674,755  
PRIOR APPLICATION NUMBER: 08/147,875A  
PRIOR FILING DATE: 1999-05-24  
NUMBER OF SEQ ID NOS: 28  
SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 18  
; LENGTH: 102  
; TYPE: PRT  
; ORGANISM: Streptococcus pneumoniae  
US-10-674-755-18

Query Match 92.5%; Score 479; DB 32; Length 102;  
Best Local Similarity 97.1%; Pred. No. 2.1e-34;  
Matches 101; Conservative 0; Mismatches 1; Indels 2; Gaps 2;

Qy	1	LAKQTELEKLLDNLDPGKTQDELDKAFAELDKKADLPNKVADLEKEISNLEITLG	60
Db	1	LAKQTELEKLLD-LDPGKTQDELDKA-AEALDKKADLPNKVADLEKEISNLEITLG	58
Qy	61	GADPEDDTALPNKLTATKAELEKTOKELDPAALNELGPDGDEE	104
Db	59	GADSEDDTALPNKLTATKAELEKTOKELDPAALNELGPDGDEE	102

Search completed: November 19, 2004, 23:17:08  
Job time : 241.024 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 19, 2004, 22:20:29 ; Search time 5.07317 Seconds  
(without alignments)  
387.238 Million cell updates/sec

Title: US-10-674-755-21

Perfect score: 518  
Sequence: 1 LAKQTELEKLDNDPEGK.....TQKEIDALNBLGPDDEE 104

Scoring table:

BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 83788 seqs, 18889656 residues

Total number of hits satisfying chosen parameters: 83788

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

Pending Patents AA New:\*  
1: /cgn2\_6/ptodata/2/paa/PCT\_NEW\_COMB.pep.\*  
2: /cgn2\_6/ptodata/2/paa/US05\_NEW\_COMB.pep.\*  
3: /cgn2\_6/ptodata/2/paa/US07\_NEW\_COMB.pep.\*  
4: /cgn2\_6/ptodata/2/paa/US08\_NEW\_COMB.pep.\*  
5: /cgn2\_6/ptodata/2/paa/US09\_NEW\_COMB.pep.\*  
6: /cgn2\_6/ptodata/2/paa/US10\_NEW\_COMB.pep.\*  
7: /cgn2\_6/ptodata/2/paa/US11\_NEW\_COMB.pep.\*  
8: /cgn2\_6/ptodata/2/paa/US60\_NEW\_COMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	127	24.5	739	US-10-732-923-3294	Sequence 3294, Ap
2	106	20.5	882	US-10-732-923-3300	Sequence 3300, Ap
3	102.5	19.8	1236	US-10-805-684-89	Sequence 89, Appl
4	101.5	19.6	281	US-10-805-684-58	Sequence 58, Appl
5	99.5	18.2	865	US-10-732-923-7041	Sequence 7041, Ap
6	97.5	18.8	1119	US-10-732-923-3321	Sequence 3321, Ap
7	95.5	18.4	880	US-10-732-923-3314	Sequence 3314, Ap
8	95.5	18.4	1146	US-10-732-923-8646	Sequence 8646, Ap
9	95.5	18.4	1178	US-10-732-923-8641	Sequence 8641, Ap
10	93.5	18.1	339	US-10-805-684-59	Sequence 59, Appl
11	93.5	18.1	864	US-10-732-923-3312	Sequence 3312, Ap
12	93	18.0	1003	US-10-732-923-3319	Sequence 3319, Ap
13	93	18.0	1047	US-10-732-923-3320	Sequence 3320, Ap
14	92.5	17.9	1960	US-10-805-684-90	Sequence 90, Appl
15	92	17.8	864	US-10-732-923-6763	Sequence 6763, Ap
16	91.5	17.7	944	US-10-732-923-3333	Sequence 3333, Ap
17	91.5	17.6	858	US-10-732-923-6800	Sequence 6800, Ap
18	90	17.4	1939	US-10-732-923-3340	Sequence 3340, Ap
19	89.5	17.3	692	US-10-732-923-3350	Sequence 3350, Ap
20	89.5	17.3	891	US-10-732-923-3347	Sequence 3347, Ap
21	89.5	17.3	1043	US-10-732-923-535	Sequence 535, Appl
22	89	17.2	811	US-10-732-923-6867	Sequence 6867, Ap
23	89	17.2	899	US-10-732-923-6873	Sequence 6873, Ap
24	89	17.2	899	US-10-732-923-6874	Sequence 6874, Ap
25	89	17.2	1903	PCT-US04-02460-3	Sequence 3, Appl

26	88.5	17.1	284	6	US-10-979-159-273	Sequence 273, Appl
27	88.5	17.1	431	6	US-10-732-923-3348	Sequence 3348, Ap
28	88	17.0	1363	6	US-10-732-923-3336	Sequence 3336, Ap
29	88	17.0	1935	6	US-10-805-684-44	Sequence 44, Appl
30	88	17.0	1939	6	US-10-805-684-93	Sequence 93, Appl
31	88	17.0	1939	6	US-10-805-684-152	Sequence 152, Appl
32	87	16.8	1127	6	US-10-732-923-3322	Sequence 3322, Ap
33	86	16.6	296	6	US-10-967-702-149	Sequence 149, Appl
34	85.5	16.5	219	6	US-10-805-684-23	Sequence 23, Appl
35	85.5	16.5	1875	6	US-10-732-923-3334	Sequence 3334, Ap
36	85.5	16.5	1875	6	US-10-732-923-3335	Sequence 3335, Ap
37	85	16.4	1085	6	US-10-732-923-3318	Sequence 3318, Ap
38	85	16.4	1805	6	US-10-732-923-3303	Sequence 3303, Ap
39	84.5	16.3	238	6	US-10-805-684-62	Sequence 62, Appl
40	84.5	16.3	691	6	US-10-955-952-16	Sequence 16, Appl
41	84.5	16.3	691	6	US-10-157-779-16	Sequence 16, Appl
42	84.5	16.3	691	6	US-10-964-241-16	Sequence 16, Appl
43	84	16.2	713	6	US-10-959-539-52	Sequence 52, Appl
44	84	16.2	811	6	US-10-732-923-6868	Sequence 6868, Ap
45	84	16.2	886	6	US-10-732-923-3310	Sequence 3310, Ap

## ALIGNMENTS

```

RESULT 1
US-10-732-923-3294
; Sequence 3294, Application US/10732923
; GENERAL INFORMATION:
; APPLICANT: Edgerton, Michael D
; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES
; FILE REFERENCE: 38-15(52796)C
; CURRENT APPLICATION NUMBER: US/10/732,923
; PRIOR APPLICATION NUMBER: 2003-12-10
; PRIOR FILING DATE: 2002-12-04
; NUMBER OF SEQ ID NOS: 24149
; SEQ ID NO 3294
; LENGTH: 739
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-10-732-923-3294

Query Match      24.5%; Score 127; DB 6; Length 739;
Best Local Similarity 34.8%; Pred. No. 0.004;
Matches 40; Conservative 15; Mismatches 40; Indels 20; Gaps 3;

QY      1 LAKQTELEKLDNDPEGKTQDELD-----KEAAEELDKKAD-----ELPNKV 45
DB      519 LEAKTALELEYVQRLEKELEIDSDSYLKEGRAPLOSKLDTKKAKLSTLEESDXT 578
      46 ADLEKEISNLEETLIGADPEEDTAA-----LPNKLATKKALEKTOKEIDALANE 95
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      579 DELDAEIKAKVQLDABGNNVNAVFKEGLEKTAAKKAEKAEKADVKKAVDE 633

RESULT 2
US-10-732-923-3300
; Sequence 3300, Application US/10732923
; GENERAL INFORMATION:
; APPLICANT: Edgerton, Michael D
; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES
; FILE REFERENCE: 38-15(52796)C
; CURRENT APPLICATION NUMBER: US/10/732,923
; PRIOR FILING DATE: 2003-12-10
; PRIOR APPLICATION NUMBER: 10/310,154
; NUMBER OF SEQ ID NOS: 24149
; SEQ ID NO 3300
; LENGTH: 882
; TYPE: PRT
; ORGANISM: Pyrococcus furiosus DSM 3638
US-10-732-923-3300

```

Query Match 20.5%; Score 106; DB 6; Length 882;

Best Local Similarity 29.1%; Pred. No. 0.036; Indels 18; Gaps 4;

Matches 34; Conservative 25; Mismatches 35; Indels 18; Gaps 4;

Qy 1 LAKQTELEKLDNDPEGKTODELDKEAAEAELDKKADELPNKVADELKEISNLEILLGA 51

Db 318 IREKLTVEELINIQ---KRISELNKESEKLEENKEKELINKLAILKEDHQLYEIK 374

Qy 52 -----ISNLSILGADPEDDTALPNKLTATKKALEKTOKELDALNEL 96

Db 375 AKKENLQLEKRLGDKSPK-DIKKLELELETKTTEBERNEIYQRIIGEL 423

## RESULT 3

US-10-805-684-89

Sequence 89, Application US/10805684

GENERAL INFORMATION:

APPLICANT: SAKAMOTO, TAKESHI

APPLICANT: TAKEDA, SHIZU

TITLE OF INVENTION: PHOS-INTERACTING PROTEINS AND USE THEREOF

FILE REFERENCE: 58748(70342)

CURRENT APPLICATION NUMBER: US/10/805,684

CURRENT FILING DATE: 2004-03-19

PRIOR APPLICATION NUMBER: 60/455,766

PRIOR FILING DATE: 2003-03-19

PRIOR APPLICATION NUMBER: 60/459,936

PRIOR FILING DATE: 2003-04-02

PRIOR APPLICATION NUMBER: 60/460,103

PRIOR FILING DATE: 2003-04-02

NUMBER OF SEQ ID NOS: 161

SOFTWARE: PatentIn Ver. 3.2

SEQ ID NO 89

LENGTH: 1236

TYPE: PRT

ORGANISM: Mus musculus

US-10-805-684-89

Query Match 19.8%; Score 102.5; DB 6; Length 1236;

Best Local Similarity 27.0%; Pred. No. 0.11; Indels 19; Gaps 3;

Matches 30; Conservative 25; Mismatches 37; Indels 19; Gaps 3;

Qy 3 KKQTELEKLDNDPEGKTODELDKEAAEAELDKKADELPNKVADELKEISNLEILLGA 62

Db 291 KQEVMSIDLEERLKKKEKTRQLEK--AKRKLDSGTTDLDQIAELQAVDELKVQLTKK 348

Qy 63 DPE-----DDTALPNKLTATKKALEKTOKELDALNELGPDDEER 104

Db 349 EEEIQLARGDDDTLHKNNAL-----KVARELQAIQIAELQEDFESEK 391

## RESULT 4

US-10-805-684-58

Sequence 58, Application US/10805684

GENERAL INFORMATION:

APPLICANT: SAKAMOTO, TAKESHI

APPLICANT: TAKEDA, SHIZU

TITLE OF INVENTION: PHOS-INTERACTING PROTEINS AND USE THEREOF

FILE REFERENCE: 58748(70342)

CURRENT APPLICATION NUMBER: US/10/805,684

CURRENT FILING DATE: 2004-03-19

PRIOR APPLICATION NUMBER: 60/455,766

PRIOR FILING DATE: 2003-03-19

PRIOR APPLICATION NUMBER: 60/459,936

PRIOR FILING DATE: 2003-04-02

PRIOR APPLICATION NUMBER: 60/460,103

PRIOR FILING DATE: 2003-04-02

NUMBER OF SEQ ID NOS: 161

SOFTWARE: PatentIn Ver. 3.2

SEQ ID NO 58

LENGTH: 281

TYPE: PRT

ORGANISM: Mus musculus

US-10-805-684-58

Query Match 19.6%; Score 101.5; DB 6; Length 281;

Best Local Similarity 27.0%; Pred. No. 0.024; Indels 19; Gaps 3;

Matches 30; Conservative 25; Mismatches 37; Indels 19; Gaps 3;

Qy 3 KKQTELEKLDNDPEGKTODELDKEAAEAELDKKADELPNKVADELKEISNLEILLGA 62

Db 165 KQEVMSIDLEERLKKKEKTRQLEK--AKRKLDSGTTDLDQIAELQAVDELKVQLTKK 222

Qy 63 DPE-----DDTALPNKLTATKKALEKTOKELDALNELGPDDEER 104

Db 223 EEEIQLARGDDDTLHKNNAL-----KVARELQAIQIAELQEDFESEK 265

## RESULT 5

US-10-732-923-7041

Sequence 7041, Application US/10732923

GENERAL INFORMATION:

APPLICANT: Edgerton, Michael D

TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES

FILE REFERENCE: 38-15(52796)C

CURRENT APPLICATION NUMBER: US/10/732,923

CURRENT FILING DATE: 2003-12-10

PRIOR APPLICATION NUMBER: 10/310,154

PRIOR FILING DATE: 2002-12-04

NUMBER OF SEQ ID NOS: 24149

SEQ ID NO 7041

LENGTH: 865

TYPE: PRT

ORGANISM: Clostridium tetani E88

US-10-732-923-7041

Query Match 19.2%; Score 99.5; DB 6; Length 865;

Best Local Similarity 30.6%; Pred. No. 0.13; Indels 27; Gaps 5;

Matches 33; Conservative 20; Mismatches 28; Indels 27; Gaps 5;

Qy 5 QTELEKLDNDPEGKT--ODELDKEAAEAELDKKADELPNKVADELKEISNLEILLGA 62

Db 406 RTEIDSLPTBLDSIKRKIFQMEIEKELAKESRSRSE--RLDELEKESLNK----- 456

Qy 63 DPEDDTAA-----LPNKLTATKKALEKTOKELDALNEL 96

Db 457 EKQEMTAKYEKEKEQIINRNILKQKLDVYKQLEKAREYD--LNKV 502

## RESULT 6

US-10-732-923-3321

Sequence 3321, Application US/10732923

GENERAL INFORMATION:

APPLICANT: Edgerton, Michael D

TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES

FILE REFERENCE: 38-15(52796)C

CURRENT APPLICATION NUMBER: US/10/732,923

CURRENT FILING DATE: 2003-12-10

PRIOR APPLICATION NUMBER: 10/310,154

PRIOR FILING DATE: 2002-12-04

NUMBER OF SEQ ID NOS: 24149

SEQ ID NO 3321

LENGTH: 1119

TYPE: PRT

ORGANISM: Daucus carota

US-10-732-923-3321

Query Match 18.8%; Score 97.5; DB 6; Length 1119;

Best Local Similarity 27.2%; Pred. No. 0.27; Indels 31; Gaps 6;

Matches 34; Conservative 28; Mismatches 32; Indels 31; Gaps 6;

Qy 2 AKQTELEKLDNDPEGKTODELDKEAAEAELDKK-----ADELPNKVADELK---EISN 54

Db 289 AREQSEIQLKLD-----BHKAILVEYKQSFEMENDKRNDFQNNVAVEVEKKEVEVK 344

Qy 55 LEIILGADPEDDTALPNK-----LATKKALEKTOKELDALNELGPD----- 99

Db 345 LEAKLAKRE-----HLDQGEKLEKEQYLASKDQDINERKSMKLEENKIDENQULL 399  
 QY 100 GDEEE 104  
 Db 400 SDOKE 404

## RESULT 7

US-10-732-923-3314  
 ; Sequence 3314, Application US/10732923  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Edgerton, Michael D  
 ; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES  
 ; FILE REFERENCE: 38-15(52796)C  
 ; CURRENT APPLICATION NUMBER: US/10/732,923  
 ; CURRENT FILING DATE: 2003-12-10  
 ; PRIOR APPLICATION NUMBER: 10/310,154  
 ; PRIOR FILING DATE: 2002-12-04  
 ; NUMBER OF SEQ ID NOS: 24149  
 ; SEQ ID NO 3314  
 ; LENGTH: 880  
 ; TYPE: PRT  
 ; ORGANISM: Pyrococcus abyssi  
 US-10-732-923-3314

Query Match 18.4%; Score 95.5; DB 6; Length 880;  
 Best Local Similarity 28.8%; Pred. No. 0.31;  
 Matches 30; Conservative 24; Mismatches 33; Indels 17; Gaps 4;

QY 1 LAKKOTLEKLDNDPEKTODELKKAEE--AEIDKKADELPNKVADL-----KXI 52  
 Db 320 LKMEBEL-KAIEVVKESGKKERAEIREFKLSIEKLEELKPYVEELEDAKQVQKI 378  
 QY 53 SNLEILGGADEPDDTALPNKLTATKKAIELEKTOKELDALNEL 96  
 Db 379 ERLKARLKGLS-----PGVETLGLSELEKERTIEBAIKI 414

## RESULT 8

US-10-732-923-8646  
 ; Sequence 8646, Application US/10732923  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Edgerton, Michael D  
 ; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES  
 ; FILE REFERENCE: 38-15(52796)C  
 ; CURRENT APPLICATION NUMBER: US/10/732,923  
 ; CURRENT FILING DATE: 2003-12-10  
 ; PRIOR APPLICATION NUMBER: 10/310,154  
 ; PRIOR FILING DATE: 2002-12-04  
 ; NUMBER OF SEQ ID NOS: 24149  
 ; SEQ ID NO 8646  
 ; LENGTH: 1146  
 ; TYPE: PRT  
 ; ORGANISM: Neurospora crassa  
 US-10-732-923-8646

Query Match 18.4%; Score 95.5; DB 6; Length 1146;  
 Best Local Similarity 30.5%; Pred. No. 0.42;  
 Matches 29; Conservative 19; Mismatches 36; Indels 11; Gaps 3;

QY 8 LEKLDNDPEKTODELKKAEEAEIDKKADELPNKVADLEKEISNLEIL-LGGADPED 66  
 Db 957 LEQENAKMEAEKETEDD-DDENDSTWVTDDDDDETVDLERAVANBEILVLGSEKPA 1015  
 QY 67 DTAALPNKLTATKKAIELEKTOKELDALNELGPGD 101  
 Db 1016 GVEVTRGEIAEKRAV-----AGAAERGBEGD 1041

RESULT 9  
 US-10-732-923-8641  
 ; Sequence 8641, Application US/10732923

; GENERAL INFORMATION:  
 ; APPLICANT: Edgerton, Michael D  
 ; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES  
 ; FILE REFERENCE: 38-15(52796)C  
 ; CURRENT APPLICATION NUMBER: US/10/732,923  
 ; CURRENT FILING DATE: 2003-12-10  
 ; PRIOR APPLICATION NUMBER: 10/310,154  
 ; PRIOR FILING DATE: 2002-12-04  
 ; NUMBER OF SEQ ID NOS: 24149  
 ; SEQ ID NO 8641  
 ; LENGTH: 1178  
 ; TYPE: PRT  
 ; ORGANISM: Neurospora crassa  
 US-10-732-923-8641

Query Match 18.4%; Score 95.5; DB 6; Length 1178;  
 Best Local Similarity 30.5%; Pred. No. 0.43;  
 Matches 29; Conservative 19; Mismatches 36; Indels 11; Gaps 3;

QY 8 LEKLDNDPEKTODELKKAEEAEIDKKADELPNKVADLEKEISNLEIL-LGGADPED 66  
 Db 989 LEQENAKMEAEKETEDD-DDENDSTWVTDDDDDETVDLERAVANBEILVLGSEKPA 1047  
 QY 67 DTAALPNKLTATKKAIELEKTOKELDALNELGPGD 101  
 Db 1048 GVEVTRGEIAEKRAV-----AGAAERGBEGD 1073

## RESULT 10

US-10-805-684-59  
 ; Sequence 59, Application US/10805684  
 ; GENERAL INFORMATION:  
 ; APPLICANT: SAKAMOTO, TAKESHI  
 ; APPLICANT: TAKEDA, SHIZU  
 ; TITLE OF INVENTION: PHOS-INTERACTING PROTEINS AND USE THEREOF  
 ; FILE REFERENCE: 58748(70342)  
 ; CURRENT APPLICATION NUMBER: US/10/805,684  
 ; CURRENT FILING DATE: 2004-03-19  
 ; PRIOR APPLICATION NUMBER: 60/455,766  
 ; PRIOR FILING DATE: 2003-03-19  
 ; PRIOR APPLICATION NUMBER: 60/459,936  
 ; PRIOR FILING DATE: 2003-04-02  
 ; PRIOR APPLICATION NUMBER: 60/460,103  
 ; PRIOR FILING DATE: 2003-04-02  
 ; NUMBER OF SEQ ID NOS: 161  
 ; SOFTWARE: PatentIn Ver. 3.2  
 ; SEQ ID NO 59  
 ; LENGTH: 339  
 ; TYPE: PRT  
 ; ORGANISM: Mus musculus  
 US-10-805-684-59

Query Match 18.1%; Score 93.5; DB 6; Length 339;  
 Best Local Similarity 27.5%; Pred. No. 0.16;  
 Matches 30; Conservative 29; Mismatches 41; Indels 9; Gaps 4;

QY 1 LAKKOTLEKLDNDP-----EGKTODELKKAEEAEIDKKADELPNKVADLEKEISNLE 56  
 Db 166 LAKLNKHEAMITDLEERLRERERKQRELEKTRKLED--STDSDQIAEQAQIAELK 223  
 QY 57 ILLGGADEPDDT--AALPNKLTATKKAIELEKTOKELDALNELGPGDDE 103  
 Db 224 MQLAKKEBSQALALARVEEAAQKMWALKI-RELETOISLEQEDLESE 271

## RESULT 11

US-10-732-923-3312  
 ; Sequence 3312, Application US/10732923  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Edgerton, Michael D  
 ; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES  
 ; FILE REFERENCE: 38-15(52796)C  
 ; CURRENT APPLICATION NUMBER: US/10/732,923

```

; CURRENT FILING DATE: 2003-12-10
; PRIOR APPLICATION NUMBER: 10/310,154
; PRIOR FILING DATE: 2002-12-04
; NUMBER OF SEQ ID NOS: 24149
; SEQ ID NO 3312
; LENGTH: 864
; TYPE: PRT
; ORGANISM: Sulfolobus solfataricus
US-10-732-923-3312

```

```

Query Match      18.1%; Score 93.5; DB 6; Length 864;
Best Local Similarity 31.9%; Pred. No. 0.46;
Matches 36; Conservative 16; Mismatches 32; Indels 29; Gaps 5;

```

```

Qy 1 LAKQTELEKLDNDPEGKTQDELDK-----AAEALDK---KA 38
Db 494 LKJNEIEIENHSHIESIKNDSEIKKINEVKELKYEEFMRLSKYTKEDLKKVKL 553
39 DELPNKVADEKEISNLEILGADPEDDTAALPNK--LATKKALEKTOKE 88
Qy 554 DEMKKKEEIEKEMRGLESELKGLDRK---ALSKLIDLKPKVKLDEMCK 602
Db

```

```

RESULT 12
US-10-732-923-3319
; Sequence 3319, Application US/10732923
; GENERAL INFORMATION:
; APPLICANT: Edgerton, Michael D
; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES
; FILE REFERENCE: 38-15(52796)C
; CURRENT FILING DATE: 2003-12-10
; PRIOR APPLICATION NUMBER: US/10/732,923
; PRIOR FILING DATE: 2003-12-10
; PRIOR APPLICATION NUMBER: 10/310,154
; NUMBER OF SEQ ID NOS: 24149
; SEQ ID NO 3319
; LENGTH: 1003
; TYPE: PRT
; ORGANISM: Kaposi's sarcoma-associated herpesvirus
US-10-732-923-3319

```

```

Query Match      18.0%; Score 93; DB 6; Length 1003;
Best Local Similarity 25.5%; Pred. No. 0.6;
Matches 27; Conservative 27; Mismatches 44; Indels 8; Gaps 2;

```

```

Qy 1 LAKQTELEKLDNDPEGKTQDELDK--EAAEALDKKDELPNKVADEKEISNLEIL 58
Db 600 LEEQOELEEQOELEEQOELEEQOELEEQOELEEQOELEEQOELEEQOELE-- 657
Qy 59 LGADPEDDTAALPNKLATKKALEKTOKEIDALNLELGPDDDEE 104
Db 658 ----EQOELEEQOELEEQOELEEQOELEEQOELEEQOELEEQOELE 699

```

```

RESULT 13
US-10-732-923-3320
; Sequence 3320, Application US/10732923
; GENERAL INFORMATION:
; APPLICANT: Edgerton, Michael D
; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES
; FILE REFERENCE: 38-15(52796)C
; CURRENT FILING DATE: 2003-12-10
; PRIOR APPLICATION NUMBER: US/10/732,923
; PRIOR FILING DATE: 2002-12-04
; NUMBER OF SEQ ID NOS: 24149
; SEQ ID NO 3320
; LENGTH: 1047
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1) ... (1047)

```

```

; OTHER INFORMATION: unsure at all Xaa locations
US-10-732-923-3320

```

```

Query Match      18.0%; Score 93; DB 6; Length 1047;
Best Local Similarity 27.5%; Pred. No. 0.63;
Matches 38; Conservative 26; Mismatches 34; Indels 40; Gaps 7;

```

```

Qy 2 AKQTELEKLDNDPEGKTQDELDKEAAEALDK-----ADELPNKVADLEK---EISN 54
Db 209 AREKVEWQKLD-----EQNTILDVKQEFVELDEKRSKSPDGLKNLVEVEKKEALITH 264
Qy 55 LE-----ILG-----GADPEDDTAALPNK-----LATKKALEKTOKE 88
Db 265 MEKVAKRQALGKKALEKLEKEIEYQKVKALREKELIKSEBKSLVTEKGIESEREB 324
Qy 89 L---DALNLELGPDDDEE 103
Db 325 LETHKAEVKIRANNEE 342

```

```

RESULT 14
US-10-805-684-90
; Sequence 90, Application US/10805684
; GENERAL INFORMATION:
; APPLICANT: SAKAMOTO, TAKESHI
; APPLICANT: TAKEDA, SHIZU
; TITLE OF INVENTION: PHOS-INTERACTING PROTEINS AND USE THEREOF
; FILE REFERENCE: 58748(70342)
; CURRENT FILING DATE: 2004-03-19
; PRIOR APPLICATION NUMBER: US/10/805,684
; PRIOR FILING DATE: 2003-03-19
; PRIOR APPLICATION NUMBER: 60/455,766
; PRIOR FILING DATE: 2003-04-02
; PRIOR APPLICATION NUMBER: 60/459,936
; PRIOR FILING DATE: 2003-04-02
; PRIOR APPLICATION NUMBER: 60/460,103
; NUMBER OF SEQ ID NOS: 161
; SOFTWARE: PatentIn Ver. 3.2
; SEQ ID NO 90
; LENGTH: 1960
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-805-684-90

```

```

Query Match      17.9%; Score 92.5; DB 6; Length 1960;
Best Local Similarity 27.0%; Pred. No. 1.4;
Matches 27; Conservative 25; Mismatches 37; Indels 11; Gaps 3;

```

```

Qy 5 QTELEKLDNDPEGKTQDELDKEAAEALDKKADLPNKVADLEKIS-----NLEI 57
Db 1503 RTEMEDLMSKDDVYKSVHELEK--SKRALQQVEEMKQLELEDELQATDPAKRLLEV 1560
Qy 58 LLGG--ADPEDDTAALPNKLATKKALEKTOKEIDALNLE 95
Db 1561 NLQAMKAQFERDLOGRDEQSEKKQLVQVREWEALIED 1600

```

```

RESULT 15
US-10-732-923-6763
; Sequence 6763, Application US/10732923
; GENERAL INFORMATION:
; APPLICANT: Edgerton, Michael D
; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES
; FILE REFERENCE: 38-15(52796)C
; CURRENT FILING DATE: 2003-12-10
; PRIOR APPLICATION NUMBER: US/10/732,923
; PRIOR FILING DATE: 2002-12-04
; NUMBER OF SEQ ID NOS: 24149
; SEQ ID NO 6763
; LENGTH: 864
; TYPE: PRT
; ORGANISM: Fusobacterium nucleatum subsp. nucleatum ATCC 25586

```

US-10-732-923-6763

Query Match	17.8%;	Score 92;	DB 6;	Length 864;
Best Local Similarity	28.1%;	Pred NO	0.62;	

Matches 34; Conservative 23; Mismatches 36; Indels 28; Gaps 5;

QY 1 LAKQTELEKLLDNLDPGKTQDE-----LDKEAEAEALDKA-----DELPNKV 45

Db 420 LTRKALQLEIKALEKE--TDDASKERLVI EKELAELENEKKVLTSMWELEKEDISKI 477

420 LTRKALQLEIEIKALEKE--TDDASKERLKVIEKELAEINKEKVLTSKMELEKEDISKI 477

QY 46 ADLEKEISNLEILLGGADPEDDTAALP---NKLATTKAELEKTQKELDALNELGPDGD 101

```

Db      478  KNKEIENVLMEKAREYDTTKLSLKYGKLTATLEKELEQGNQND-----KDGK 530
          ::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

```

QY 102 E 102

Db 531 E 531

Search completed: November 19, 2004, 23:18:04  
Job time : 6.07317 secs

Job time : 6.07317 secs

**This Page Blank (uspto)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using SW model

Run on: November 19, 2004, 22:05:43 ; Search time 13.9512 Seconds  
(without alignments)  
717.252 Million cell updates/sec

Title: US-10-674-755-21

Perfect score: 518  
Sequence: 1 LAKKQTELEKLDNDPEBK.....TKELDAALNELGPDDEBE 104

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 79:\*

1: PIR1:\*\n2: PIR2:\*\n3: PIR3:\*\n4: PIR4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	484	93.4	744	2 P95013	pneumococcal surfa
2	128	24.7	619	2 A97887	surface protein ps
3	128	24.7	619	2 A41971	myosin heavy chain
4	109.5	21.1	924	2 S06117	myosin heavy chain
5	109.5	21.1	2007	1 B43402	hypothetical prote
6	104	20.1	281	2 F75216	hypothetical prote
7	103.5	20.0	388	2 A46173	Mtp4 protein - Str
8	103.5	20.0	1938	1 A40997	myosin heavy chain
9	103.5	20.0	1976	2 A59252	myosin heavy chain
10	103	19.9	397	2 H86754	prophage p12 prote
11	101	19.5	2139	2 T18296	myosin heavy chain
12	100.5	19.4	537	2 S33068	myosin heavy chain
13	100.5	19.4	1509	1 A27224	myosin heavy chain
14	100.5	19.4	1940	2 A59287	myosin heavy chain
15	100	19.3	259	2 D60110	repetitive protein
16	100	19.3	385	2 T20410	hypothetical prote
17	100	19.3	1837	2 T41023	probable nuclear p
18	99.5	19.2	1992	2 A47297	myosin heavy chain
19	99	19.1	199	2 A61513	myosin, sarcomeric
20	99	19.1	388	2 S52536	fcra 15 protein -
21	98.5	19.0	1964	2 A59282	nonmuscle myosin I
22	98	18.9	415	2 S35760	myosin heavy chain
23	97.5	18.8	1119	2 T14321	nuclear matrix con
24	97.5	18.8	1972	1 A41604	myosin heavy chain
25	97.5	18.8	2346	2 T13829	Tpr homolog - fru1
26	97	18.7	518	2 G84488	En/gpm-like transp
27	97	18.7	587	2 JCL119	Fc gamma (1g) rec
28	97	18.7	936	2 S39083	myosin heavy chain
29	97	18.7	1957	2 A59294	skeletal myosin -

30	96.5	18.6	405	2 A33939	Fc gamma (1g) rec
31	96.5	18.6	629	2 T44607	hypothetical prote
32	96	18.5	1169	2 A64505	p15 homolog Met
33	96	18.5	1175	2 C35815	myosin heavy chain
34	96	18.5	1175	2 D35815	myosin heavy chain
35	96	18.5	1201	2 A35815	myosin heavy chain
36	96	18.5	1201	2 B35815	myosin heavy chain
37	96	18.5	1313	2 A48467	myosin heavy chain
38	96	18.5	2385	2 A32491	myosin heavy chain
39	96	18.5	2411	2 B32491	myosin heavy chain
40	95.5	18.4	880	2 F75103	conserved hypotet
41	95.5	18.4	1367	2 F84669	probable chromosom
42	95.5	18.4	1937	2 I38055	myosin heavy chain
43	95.5	18.4	1938	2 JCS421	smooth muscle myos
44	95.5	18.4	1972	2 JCS420	smooth muscle myos
45	95	18.3	516	2 B84709	hypothetical prote

#### ALIGNMENTS

RESULT 1  
P95013  
pneumococcal surface protein A [imported] - Streptococcus pneumoniae (strain TIGR4)  
C:Species: Streptococcus pneumoniae  
C:Date: 03-Aug-2001 #sequence\_revision 03-Aug-2001 #ext\_change 09-Jul-2004  
C:Accession: P95013  
R:Teitelin, H.; Neilson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; Hei-  
on, J.D.; Umayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzapple,  
nson, T.; Hickey, B.K.; Holt, I.E.  
Science 293, 498-506, 2001  
A:Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrison  
A>Title: Complete Genome Sequence of a Virulent Isolate of Streptococcus pneumoniae.  
A:Reference number: A95000; MUID:21357209; PMID:11463916  
A:Accession: P95013  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-744 <KUR>  
A:Cross-references: UNIPROT:Q97739; GB:AEO05672; PIDN:AAK74303.1; PID:G14971584; GSPDB:  
A:Experimental source: strain TIGR4  
C:Genetics:  
A:Gene: SP0117

Query Match 93.4%; Score 484; DB 2; Length 744;  
Best Local Similarity 95.2%; Pred. No. 2.9e-26;  
Matches 99; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 LAKKQTELEKLDNDPEBKTDDEKAEAEELDKKADLPNKVADLEKEISNLEILG 60  
DB 346 LAKKQTELEKLDNDPEBKTDDEKAEAEELDKKADLPNKVADLEKEISNLEILG 405

QY 61 GADPEDDTAALPNKATKKALEKTKQKELDAALNELGPDDEBE 104  
DB 406 GADSEDDTAALQNKATKKALEKTKQKELDAALNELGPDDEBE 449

RESULT 2  
A97887  
surface protein pspA precursor [imported] - Streptococcus pneumoniae (strain R6)  
C:Species: Streptococcus pneumoniae  
C:Date: 22-Oct-2001 #sequence\_revision 22-Oct-2001 #ext\_change 09-Jul-2004  
C:Accession: A97887  
R:Hoskins, J.A.; Alborn Jr., W.; Arnold, J.; Blaszcak, L.; Burgett, S.; Dehoff, B.S.; E-  
e, R.; Leblanc, D.J.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Matsushima, P.; McAhren, S.; I-  
Y, P.; Sun, P.M.; Winkler, M.E.  
J. Bacteriol. 183, 5709-5717, 2001  
A:Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaekunas, S.R.,  
A>Title: Genome of the Bacterium Streptococcus pneumoniae Strain R6.  
A:Reference number: A97887; MUID:21429245; PMID:11544234  
A:Accession: A97887  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-619 <KUR>

A:Cross-references: UNIPROT:Q54972; UNIPROT:Q8DR10; GB:AE007317; PIDN:AAK98925.1; PID:94  
C:Genetics:  
A:Gene: pspa

Query Match 24.7%; Score 128; DB 2; Length 619;  
Best Local Similarity 34.8%; Pred. No. 0.09;  
Matches 39; Conservative 15; Mismatches 38; Indels 20; Gaps 3;

Qy 4 KQTELEKLDNLDEGKTQDELD-----KEAAEALDKKAD-----ELPNKAVDL 48  
Db 208 KIALENOVHRLLEQLEKIDSESDYAKGEFRAPLOSKLDAAKAKLSKEELSDKIDEL 267

Qy 49 EKEISNLEILLGADPEDDT-----AALPNKLATTKALEKTQKELDAALNE 95  
Db 268 DAELAKLEDQKAAEENNNDYDFKEGLEKTIKAKKALEKTEADLKAVNE 319

RESULT 3  
A41971  
surface protein pspa precursor - Streptococcus pneumoniae  
N:Alternate names: pneumococcal surface protein A  
C:Species: Streptococcus pneumoniae  
C:Date: 04-Mar-1993 #sequence\_revision 18-Nov-1994 #text\_change 09-Jul-2004  
C:Accession: A41971; A60282; A31314  
R:Yoher, J.; Briles, D.E.  
J. Bacteriol. 174, 601-609, 1992  
A:Title: Structural properties and evolutionary relationships of Pspa, a surface protein  
A:Reference number: A41971; MUID:92105030; PMID:11729249  
A:Accession: A41971  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-619 <YOT>  
A:Cross-references: UNIPROT:Q54972; UNIPROT:Q8DR10; GB:M74122; NID:G153840; PIDN:AAA2701  
A:Note: sequence extracted from NCBI backbone (NCBI:75635, NCBI:75636)  
R:Talkington, D.F.; Crimmins, D.L.; Voellinger, D.C.; Yoher, J.; Briles, D.E.  
Infect. Immun. 59, 1285-1289, 1991  
A:Title: A 43-kilodalton pneumococcal surface protein, Pspa: isolation, protective ability  
A:Reference number: A60282; MUID:91169596; PMID:2004810  
A:Accession: A60282  
A:Molecule type: protein  
A:Residues: 32-76 <TAL>  
A:Experimental source: strain JY2008  
C:Genetics:  
A:Gene: pspa  
F:1-31/Domain: signal sequence #status predicted <SIG>  
F:32-61/Domain: surface protein pspa #status predicted <MAT>  
F:411-430/Domain: cpl repeat homology <CP01>  
F:431-450/Domain: cpl repeat homology <CP02>  
F:451-470/Domain: cpl repeat homology <CP03>  
F:471-490/Domain: cpl repeat homology <CP04>  
F:491-510/Domain: cpl repeat homology <CP05>  
F:511-530/Domain: cpl repeat homology <CP06>  
F:531-550/Domain: cpl repeat homology <CP07>  
F:551-570/Domain: cpl repeat homology <CP08>  
F:571-591/Domain: cpl repeat homology <CP09>  
F:592-611/Domain: cpl repeat homology <CP10>

Query Match 24.7%; Score 128; DB 2; Length 619;  
Best Local Similarity 34.8%; Pred. No. 0.09;  
Matches 39; Conservative 15; Mismatches 38; Indels 20; Gaps 3;

Qy 4 KQTELEKLDNLDEGKTQDELD-----KEAAEALDKKAD-----ELPNKAVDL 48  
Db 208 KIALENOVHRLLEQLEKIDSESDYAKGEFRAPLOSKLDAAKAKLSKEELSDKIDEL 267

Qy 49 EKEISNLEILLGADPEDDT-----AALPNKLATTKALEKTQKELDAALNE 95  
Db 268 DAELAKLEDQKAAEENNNDYDFKEGLEKTIKAKKALEKTEADLKAVNE 319

RESULT 4  
S06117  
myosin heavy chain, nonmuscle (clone lambda-FMHC) - chicken (fragment)

C:Species: Gallus gallus (chicken)  
C:Date: 30-Sep-1991 #sequence\_revision 30-Sep-1991 #text\_change 09-Jul-2004  
C:Accession: S06117  
R:Katsuragawa, Y.; Yanagisawa, M.; Inoue, A.; Masaki, T.  
Eur. J. Biochem. 184, 611-616, 1989  
A:Title: Two distinct nonmuscle myosin-heavy-chain mRNAs are differentially expressed in  
s.  
A:Reference number: S06116; MUID:90032648; PMID:2806244  
A:Accession: S06117  
A:Status: not compared with conceptual translation  
A:Molecule type: mRNA  
A:Residues: 1-924 <KAT>  
A:Cross-references: UNIPROT:Q02015; GB:X17590  
A:Note: this translation is not annotated in GenBank entry GGMHCFC, release 114  
C:Superfamily: myosin heavy chain; myosin motor domain homology  
F:1-303/Domain: myosin motor domain homology (fragment) <MOT>

Query Match 21.1%; Score 109.5; DB 2; Length 924;  
Best Local Similarity 31.2%; Pred. No. 2.6;  
Matches 34; Conservative 27; Mismatches 41; Indels 7; Gaps 3;

Qy 1 LAKKQTELEKLDNLDE-----PEGKTQDELDKEAAEALDKKADLPKRVADLEKISNLE 56  
Db 557 LAKLKNQEMWITDLERLKEKERTQELK--AKRKLDSSTTDLQQLALQELQLELK 614

Qy 57 ILLGADPEDDTA-ALPNKLATTKALEKTQKELDAALNEGPDGDEE 104  
Db 615 IQLAKKEELQALALRGDEEAVQNNALKYIRELQALQALQELQLEBK 663

RESULT 5  
B43402  
myosin heavy chain-B, neuronal - chicken  
N:Contains: myosin ATPase (EC 3.6.4.1)  
C:Species: Gallus gallus (chicken)  
C:Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 09-Jul-2004  
C:Accession: B43402; A43402  
R:Takahashi, M.; Kawamoto, S.; Adelstein, R.S.  
J. Biol. Chem. 267, 17864-17871, 1992  
A:Title: Evidence for inserted sequences in the head region of nonmuscle myosin specific  
yosin.  
A:Reference number: A43402; MUID:92388144; PMID:1355479  
A:Accession: B43402  
A:Molecule type: mRNA  
A:Residues: 1-2007 <TAX>  
A:Cross-references: UNIPROT:Q02015; GB:M93676; NID:G212448; PIDN:AAA48988.1; PID:G212452  
A:Note: the sequence of residues 212-221 and 632-652 and the corresponding nucleotide seq  
A:Accession: A43402  
A:Molecule type: mRNA  
A:Residues: 1-211;222-631;653-2007 <TA2>  
A:Cross-references: GB:M93676; NID:G212448; PIDN:AAA48985.1; PID:G212449  
A:Note: sequence extracted from NCBI backbone (NCBI:112664)  
C:Comment: Alternatively spliced segments 1 and 2 are found exclusively in nonmuscle myot  
C:Keywords: actin binding; myosin heavy chain; myosin motor domain homology  
F:1-2007/Product: myosin heavy chain-B, neuronal #status predicted <MYN>  
F:1-211;222-631;653-2007/Product: myosin heavy chain-B, nonmuscle #status predicted <MYS>  
F:88-802/Domain: myosin motor domain homology <MOT>  
F:178-185/Region: nucleotide-binding motif A (P-loop)  
F:212-221/Region: alternatively spliced segment 1 #status experimental  
F:559-593/Region: actin binding #status predicted  
F:632-652/Region: alternatively spliced segment 2 #status experimental  
F:659-714/Region: actin binding #status predicted  
F:875-2007/Domain: coiled coil #status predicted <COI>  
F:875-1315/Region: S2  
F:1316-2007/Region: light meromyosin  
F:129/Modified site: N6, N6, N6-trimethyllysine (Lys) #status predicted  
F:184/Binding site: ATP (Lys) #status predicted  
F:732, 742/Active site: Cys #status predicted  
F:1954/Binding site: phosphate (Thr) (covalent) #status predicted  
F:1987/Binding site: phosphate (Ser) (covalent) #status predicted

Query Match 21.1%; Score 109.5; DB 1; Length 2007;

Best Local Similarity 31.2%; Pred. No. 5.8;  
Matches 34; Conservative 27; Mismatches 41; Indels 7; Gaps 3;

QY 1 LAKKQTELEKLDNDL-----PEKGTODELDKKAALAEALDKKADLPNKVADLEKEISNLE 56  
Db 1056 LAKKQTELEKLDNDL-----PEKGTODELDKKAALAEALDKKADLPNKVADLEKEISNLE 1113  
QY 57 ILLGADPEDDTA-ALPNKLTATKKALEKTEKELDAALNLEGPDDGDEE 104  
Db 1114 ILLGADPEDDTA-ALPNKLTATKKALEKTEKELDAALNLEGPDDGDEE 1162

## RESULT 6

F75216  
hypothetical protein PAB2181 - Pyrococcus abyssi (strain Orya)

C:Species: Pyrococcus abyssi  
C:Date: 20-Aug-1999 #sequence\_revision 20-Aug-1999 #text\_change 09-Jul-2004  
C:Accession: F75216  
R:anonymous, Genoscope  
A:Description: Pyrococcus abyssi genome sequence: insights into archaeal chromosome stru  
A:Reference number: A75001  
A:Accession: F75216  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-281 <KAW>  
A:Cross-references: UNIPROT:Q9V217; GB:AJ248283; GB:AL096836; NID:95457433; PIDN:CA84918  
A:Experimental source: strain Orya  
C:Gene: PAB2181

Query Match 20.1%; Score 104; DB 2; Length 281;  
Best Local Similarity 34.2%; Pred. No. 1.8;  
Matches 25; Conservative 17; Mismatches 19; Indels 12; Gaps 1;

QY 2 AKKQTELEKLDNDLPEKGTODELDKKAALAEALDKKADLPNKVADLE 49  
Db 200 AKKQTELEKLDNDLPEKGTODELDKKAALAEALDKKADLPNKVADLE 259  
QY 50 KEISNLEILLGGA 62  
Db 260 KEISNLEILLGGA 272

## RESULT 7

A46173  
Mip4 protein - Streptococcus sp. (group A)

C:Species: Streptococcus sp.  
C:Date: 21-Sep-1993 #sequence\_revision 25-Apr-1997 #text\_change 30-May-1997  
C:Accession: A46173  
R:O'Tool, P.; Stenberg, L.; Riseler, M.; Lindahl, G.  
Proc. Natl. Acad. Sci. U.S.A. 89, 8661-8665, 1992  
A:Title: Two major classes in the M protein family in group A streptococci.  
A:Reference number: A46173; MUID:92409576; PMID:1528877  
A:Contents: group A  
A:Accession: A46173  
A:Status: preliminary  
A:Molecule type: nucleic acid  
A:Residues: 1-388 <OLT>  
A>Note: sequence extracted from NCBI backbone (NCBIN:114063, NCBIPI:114064)  
C:Superfamily: M5 protein

Query Match 20.0%; Score 103.5; DB 2; Length 388;  
Best Local Similarity 32.1%; Pred. No. 2.7;  
Matches 35; Conservative 24; Mismatches 43; Indels 7; Gaps 4;

QY 1 LAKKQTELEKLDNDLPEKGTODELDKKAALAEALDKKADLPNKVADLEKEISNLE 55  
Db 162 LAKKQTELEKLDNDLPEKGTODELDKKAALAEALDKKADLPNKVADLEKEISNLE 241  
QY 56 ILLGADPEDDTAALPNKLTATKKALEKTEKELDAALNLEGPDDGDEE 104  
Db 242 ILLGADPEDDTAALPNKLTATKKALEKTEKELDAALNLEGPDDGDEE 288

RESULT 8:  
A40997  
myosin heavy chain, striated adductor muscle - scalllop (Aequipecten irradians)

N:Contains: myosin ATPase (EC 3.6.4.1)  
C:Species: Aequipecten irradians  
C:Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 09-Jul-2004  
C:Accession: A40997; S13557  
R:Mytilus, L.; Goodwin, E.B.; Szent-Gyorgyi, A.G.  
J. Biol. Chem. 266, 18469-18476, 1991

A:Reference number: A40997; MUID:92011595; PMID:1917970  
A:Accession: A40997  
A:Molecule type: mRNA  
A:Residues: 1-1938 <NYI>  
A:Cross-references: UNIPROT:P24733; GB:X55714; NID:95611; PIDN:CAA39247.1; PID:95612  
C:Superfamily: myosin heavy chain; myosin motor domain homology  
C:Keywords: actin binding; ATP; coiled coil; hydrophobic; muscle contraction; nucleotide ;  
F:176-183/Region: myosin motor domain homology <MMOT>  
F:194-201/Region: myosin motor domain homology  
F:547-586/Region: actin binding #status predicted  
F:653-675/Region: actin binding #status predicted  
F:836-1938/Domain: coiled coil #status predicted <COI>  
F:836-1276/Region: S2  
F:1277-1938/Region: light meromyosin  
F:182/Binding site: ATP (lys) #status predicted  
F:693,703/Active site: Cys #status predicted

Query Match 20.0%; Score 103.5; DB 1; Length 1938;  
Best Local Similarity 27.7%; Pred. No. 15;  
Matches 36; Conservative 21; Mismatches 42; Indels 31; Gaps 3;

QY 2 AKKQTELEKLDNDLPEKGTODELDKKAALAEALDKKADLPNKVADLE 49  
Db 1022 AKKQTELEKLDNDLPEKGTODELDKKAALAEALDKKADLPNKVADLE 1081  
QY 50 KEISNLEILLGGADEDDTAALPNKLTATKKALEKTEKELDA 92  
Db 1082 KEISNLEILLGGADEDDTAALPNKLTATKKALEKTEKELDA 1139

QY 93 INELGPDGDE 102  
Db 1140 INELGPDGDE 1149

## RESULT 9

A59252

myosin heavy chain, nonmuscle, form IIB - human

N:Alternate names: myosin type 10; NMHC-B  
N:Contains: myosin ATPase (EC 3.6.4.1)  
C:Species: Homo sapiens (man)  
C:Date: 19-May-2000 #sequence\_revision 19-May-2000 #text\_change 09-Jul-2004  
C:Accession: A59252; B61231; G02055  
R:Phillips, C.L.; Yamakawa, K.; Adelstein, R.S.  
J. Muscle Res. Cell. Motil. 16, 379-389, 1995  
A:Title: Cloning of the cDNA encoding human nonmuscle myosin heavy chain-B and analysis  
A:Reference number: A59252; MUID:96025307; PMID:7499478  
A:Accession: A59252

A:Status: not compared with conceptual translation

A:Molecule type: mRNA

A:Residues: 1-1976 <STM>

A:Cross-references: UNIPROT:P35580; GB:M69181; NID:9641957; PIDN:AAA99177.1; PID:9641958

A:Experimental source: clone 11b lambda Zap II adult human T-cell library; cell line Jur

A>Note: between nucleotides 1942-1943 in mRNA encoding human brain HSC-B there is an alt

R:Simons, M.; Wang, M.; McBride, O.W.; Kawamoto, S.; Yamakawa, K.; Gdula, D.; Adelstein,

Circ. Res. 69, 530-539, 1991

A:Title: Human nonmuscle myosin heavy chains are encoded by two genes located on differe

A:Reference number: A61231; MUID:91316803; PMID:1860190

A:Accession: B61231

A:Molecule type: mRNA

A:Residues: 63-237 'K', 239-664 'L', 666-722 <ST2>

A:Cross-references: GB:M69181; NID:9641957

R:Meir, L.  
submitted to the EMBL Data Library, August 1995  
A:Reference number: H00753  
A:Accession: G02055  
A:Status: translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-81 <MEI>  
A:Cross-references: EMBL:U34304; NID:G143217; PIDN:AAA84880.1; PID:G143218  
C:Genetics:  
A:Gene: GDB:MYH10  
A:Cross-references: GDB:127350; GDB:G00-127-350; OMIM:160776  
A:Map position: 17p13-17p13  
C:Superfamily: myosin heavy chain; myosin motor domain homology  
F:88-771/Domain: actin binding; ATP; coiled coil; hydrolyase; methylated amino acid; nucleotide  
F:178-185/Region: nucleotide-binding motif A (P-loop)  
F:559-572/Region: actin binding #status predicted  
F:633-647/Region: actin binding #status predicted  
F:129/Modified site: N6, N6, N6-trimethyllysine (Lys) #status predicted  
F:184/Binding site: ATP (Lys) #status predicted  
F:701,711/Active site: Cys #status predicted

Query Match 20.0%; Score 103.5; DB 2; Length 1976;  
Best Local Similarity 27.9%; Pred. No. 15;  
Matches 31; Conservative 24; Mismatches 37; Indels 19; Gaps 3;

OY 3 KKQTELEKLDNLDPEGKTQDELKKAELDKKADLPNKVADLEKEISNLEILGGA 62  
DB 1031 KQEWISDLEERLKEETKROELK--AKRKLDSITTDLQOQIALQDELKQLAKK 1088

OY 63 DPE-----DDTALPNKLTATKALEKTOKELDALNELGPDGDEER 104  
DB 1089 EEEQLGALARGDDFTLHKNNAL-----KVRLELQALVELQEDFESK 1131

RESULT 10  
H86754  
prophage p12 protein 33, capsid protein [imported] - Lactococcus lactis subsp. lactis (B  
C:Species: Lactococcus lactis subsp. lactis  
C:Date: 23-Mar-2001 #sequence\_revision 23-Mar-2001 #text\_change 09-Jul-2004  
C:Accession: H86754  
R:Boletín, A.; Wincker, P.; Manger, S.; Jallón, O.; Malarme, K.; Weisenbach, J.; Ehrlich  
Genome Res. 11, 731-753, 2001  
A:Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis ss  
A:Reference number: A86825; MUID:21235186; PMID:11337471  
A:Accession: H86754  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-397 <STO>  
A:Cross-references: UNIPROT:Q9CGQ6; GB:AE005176; PID:G12723992; PIDN:AAK05138.1; GSPDB:C  
A:Experimental source: strain ILL403  
C:Genetics:  
A:Gene: p1233

Query Match 19.9%; Score 103; DB 2; Length 397;  
Best Local Similarity 34.1%; Pred. No. 3;  
Matches 31; Conservative 20; Mismatches 32; Indels 8; Gaps 4;

OY 1 LAKQTELEKLDNLDPEGKTQDELKKAELDKKADLPNKVADLEKEISNLE-IL 58  
DB 29 LEKQNDLEERALE---EAKTDEISITVSDSADLEKQVLDLEKIALQKEXODLEDEL 84

OY 59 LGADPEDDTALPNKLTATKALEKTOKELE 89  
DB 85 AKAADPTDQKPKDSEKRRKKFKV--TEEL 113

RESULT 11  
T18296  
myosin heavy chain - Entamoeba histolytica  
C:Species: Entamoeba histolytica  
C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004  
C:Accession: T18296

R:Guillen, N.  
submitted to the EMBL Data Library, February 1997  
A:Reference number: Z18865  
A:Accession: T18296  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-2139 <GUI>  
A:Cross-references: UNIPROT:Q07569; EMBL:L03534; NID:G1850912; PID:G1850913; PIDN:AA84806  
C:Genetics:  
A:Gene: nhcA  
C:Superfamily: myosin heavy chain; myosin motor domain homology  
F:91-780/Domain: myosin motor domain homology <MMO>

Query Match 19.5%; Score 101; DB 2; Length 2139;  
Best Local Similarity 30.1%; Pred. No. 24;  
Matches 37; Conservative 23; Mismatches 33; Indels 30; Gaps 6;

OY 2 AKQTELEKLDNLDPEGKTQDEL-----DKAAAEELDK--KADELPNKV 45  
DB 1485 AKQSDAD--LEELNKTVEEHDEVAKLNTQITKLRDNOQSAEELNELSKADKDKKI 1542

OY 46 ADLEKEISNLE--ILLGADP-----EDDTALPNKLTATKKA---ELEKTOKELDAL 93  
DB 1543 SELEEQVELBSRPVGTGNADENEKIKRDAQIADLNPALEKGVQNNQLQATNKELAKD 1602

OY 94 NEL 96  
DB 1603 NDL 1605

RESULT 12  
S33068  
myosin heavy chain - fluke (Schistosoma mansoni) (fragment)  
N:Alternate names: surface antigen, 200X  
C:Species: Schistosoma mansoni  
C:Date: 22-Nov-1993 #sequence\_revision 06-Sep-1996 #text\_change 09-Jul-2004  
C:Accession: S33068  
R:Soisson, L.M.A.; Masters, C.P.; Tom, T.D.; McNally, M.T.; Lowell, G.H.; Strand, M.  
J. Immunol. 149, 3612-3620, 1992  
A:Title: Induction of protective immunity in mice using a 62-kDa recombinant fragment of  
A:Reference number: A46514; MUID:93056536; PMID:1431131  
A:Accession: S33068  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-527 <SOI>  
A:Cross-references: UNIPROT:Q26589; EMBL:X65591  
A:Note: the authors translated the codon CAA for residue 346 as Lys  
C:Superfamily: myosin heavy chain; myosin motor domain homology  
C:Keywords: ATP; surface antigen

Query Match 19.4%; Score 100.5; DB 2; Length 527;  
Best Local Similarity 28.4%; Pred. No. 6.1;  
Matches 31; Conservative 23; Mismatches 40; Indels 15; Gaps 3;

OY 2 AKQTELEKLDNLDPEGKTQDELKKAELDKK-----DELPNKVADLEKEISNLE 56  
DB 308 AKLESTIDEMENLARKQKIRGVDEKSKRKLGDKLQETVDDLEKVDLEQLRKE 367

OY 57 ILLGADP--EDDTALPNKLTATKALEKTOKELDALNELGPDGDEE 103  
DB 368 AEIGLSGKFEDEGLV-----ADLQRRKIKELQTRIOLEEDLEAE 408

RESULT 13  
A27224  
myosin heavy chain II - Acanthamoeba castellanii  
N:Contains: myosin ATPase (EC 3.6.4.1)  
C:Species: Acanthamoeba castellanii  
C:Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 09-Jul-2004  
C:Accession: A27224  
R:Hammer III, J.A.; Bowers, B.; Paterson, B.M.; Korn, E.D.  
J. Cell Biol. 105, 913-925, 1987  
A:Title: Complete nucleotide sequence and deduced polypeptide sequence of a nonmuscle myr  
A:Reference number: A27224; MUID:87308395; PMID:3040773

A/Accession: A27224  
 A/Molecule type: DNA  
 A/Residues: 1-1509 <HAM>  
 A/Cross-references: UNIPROT:P05655; GB:Y00624; GB:M12702; GB:M12703; GB:M19549; NID:9558  
 C/Genetics:  
 A/Intons: 69/3; 119/3; 181/2  
 C/Superfamily: myosin heavy chain; myosin motor domain homology  
 C/Keywords: actin binding; ATP; coiled coil; hydrolase; methylated amino acid; nucleotide  
 F:92-775/Domain: myosin motor domain homology <MMOT>  
 F:182-189/Region: nucleotide-binding motif A (P-loop)  
 F:544-576/Region: actin binding #status predicted  
 F:660-682/Region: actin binding #status predicted  
 F:848-1227/Domain: coiled coil #status predicted <COI>  
 F:1228-1247/Domain: hinge <HIN>  
 F:1248-1482/Domain: coiled coil #status predicted <CO2>  
 F:1483-1509/Domain: carboxyl-terminal <CBT>  
 F:133/Modified site: N6,N6,N6-trimethyllysine (lys) #status predicted  
 F:188/Binding site: ATP (lys) #status predicted

Query Match 19.4%; Score 100.5; DB 1; Length 1509;  
 Best Local Similarity 31.5%; Pred. No. 18;  
 Matches 39; Conservative 23; Mismatches 31; Indels 31; Gaps 6;

QY 5 QTELEKLDNDPEGKTODELDKEA-----AEALDKK-----ADELPNKVADLEKEI 52  
 DB 851 QNRFOKEIDDLKKQVK---DLEKELAAKQANAKLDKEKQLAEDADKLEKDLALKTKI 907

QY 53 SNEILLGG--ADPEDDTAALPNKLTATKKALE-----KTQKELDALNELGPDG 100  
 DB 908 LDLE----GEKADLEEDNALQKKVAGLEBELQETSASNDILEQRKLEAEKGELEKASL 963

QY 101 DEEE 104  
 DB 964 EEE 967

RESULT 14  
 A59287  
 myosin heavy chain - fluke (Schistosoma mansoni) (strain Brazilian LE)  
 C/Species: Schistosoma mansoni  
 C/Date: 09-Jun-2000 #sequence\_revision 09-Jun-2000 #text\_change 09-Jul-2004  
 C/Accession: A59287  
 R/Weston, D.S.; Schmitz, J.; Kemp, M.; Kunz, W.  
 M.O. Blochem. Parasitol. 58, 161-164, 1993  
 A/Title: Cloning and sequence characterization of a complete myosin heavy chain cDNA fro  
 A/Reference number: A59287; MUID:93211444; PMID:8459827  
 A/Accession: A59287  
 A/Status: preliminary; not compared with conceptual translation  
 A/Molecule type: mRNA  
 A/Residues: 1-1940 <MES>  
 A/Cross-references: UNIPROT:Q02456; GB:L01634; PIDN:AAA29905.1  
 A/Experimental source: strain Brazilian LE  
 C/Genetics:  
 A/Gene: MYH  
 C/Superfamily: myosin heavy chain; myosin motor domain homology  
 F:82-752/Domain: myosin motor domain homology <MMO>

Query Match 19.4%; Score 100.5; DB 2; Length 1940;  
 Best Local Similarity 28.4%; Pred. No. 23;  
 Matches 31; Conservative 23; Mismatches 40; Indels 15; Gaps 3;

QY 2 AKQTELEKLDNDPEGKTODELDKEAAEALELDKA-----DELPNKVADLEKEISNLE 56  
 DB 1011 AKLESTLDEEENLAEKQKRGVEKSKRLTEGLDKATQETVDDLERYKRDLEELRRKE 1070

QY 57 ILLGGADP--EDDTAALPNKLTATKKALEKTOKELDALNELGPDGDE 103  
 DB 1071 AEIGSGSKPEDEQIV-----AQLQKIKELQTRIQLEEDLEAE 1111

RESULT 15  
 D60110  
 repetitive protein antigen 3 - Trypanosoma cruzi (fragment)

C/Species: Trypanosoma cruzi  
 C/Date: 10-Nov-1992 #sequence\_revision 10-Nov-1992 #text\_change 09-Jul-2004  
 C/Accession: D60110  
 R/Hofc, D.F.; Kim, K.S.; Otsu, K.; Moser, D.R.; Yost, W.J.; Blumlin, J.H.; Donelson, J.E  
 Infect. Immun. 57, 1959-1967, 1989  
 A/Title: Trypanosoma cruzi expresses diverse repetitive protein antigens.  
 A/Reference number: A60110; MUID:89277508; PMID:2659529  
 A/Accession: D60110  
 A/Status: not compared with conceptual translation  
 A/Molecule type: mRNA  
 A/Residues: 1-259 <HOF>  
 A/Cross-references: UNIPROT:Q7M3R6  
 C/Keywords: tandem repeat  
 F:2-99/Region: 14-residue repeats

Query Match 19.3%; Score 100; DB 2; Length 259;  
 Best Local Similarity 34.3%; Pred. No. 3.1;  
 Matches 34; Conservative 16; Mismatches 35; Indels 14; Gaps 4;

QY 3 KQTELEKLDNDPEG---KTQDELDEAAEALELDKKADELPMNKVADLEKEISNLEIL 58  
 DB 2 QKAAENRLEADELEQKAAENRLEADELEQKTAENE--RLADELEQKAAENRLEADELE-- 57

QY 59 LGGADPEDDTAALPNKLTATKKALEKTOKELD--AALNE 95  
 DB 58 ----QKAAENRLEADELEQKAAENRLEADELEQKAAENE 92

Search completed: November 19, 2004, 22:36:10  
 Job time : 14.9512 secs

**This Page Blank (uspto)**

DE PSPA (Frag

Psya (Fragment).

```
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCB1_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=121;
RX MEDLINE=20472698; PubMed=11015380;
RA Beall B., Gherardi G., Facklam R.R., Hollingshead S.K.;
RT "Pneumococcal pspa sequence types of prevalent multiresistant
RT pneumococcal strains in the United States and of internationally
RT disseminated clones.";
RL J. Clin. Microbiol. 38:3663-3669(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=121;
RX Beall B.W.;
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF255547; AAF68100.1; -.
FT NON TER 1 1
SQ SEQUENCE 241 AA; 26038 MW; BB87E1AC25FA669 CRC64;

Query Match 96.9%; Score 502; DB 2; Length 241;
Best Local Similarity 98.1%; Pred. No. 3.5e-26;
Matches 102; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Cy 1 LAKQTELEKLDLNDPEGKTODELDKKAABAEIDKKADELPNKVADLEKEISNLEITLG 60
Db 77 LAKQTELEKLDLNDPEGKTODELDKKAABAEIDKKADELPNKVADLEKEISNLEITLG 136

Cy 61 GADPEDDTAALPNKLTATKKALEKTKOKELDALNELGPDGDEE 104
Db 137 GADPEDDTAALQNKLTATKKALEKTKOKELDALNELGPDGDEE 180

RESULT 3
ID Q9LAX8 PRELIMINARY; PRT; 502 AA.
AC Q9LAX8;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Pspa (Fragment).
GN Name=pspa;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCB1_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BG8090;
RX MEDLINE=20448953; PubMed=10992499;
RA Hollingshead S.K., Becker R., Briles D.E.;
RT "Diversity of Pspa: mosaic genes and evidence for past recombination
RT in Streptococcus pneumoniae.";
RL Infect. Immun. 68:5889-5900(2000).
DR EMBL; AF071817; AAF27713.1; -.
DR InterPro; IPR000533; Tropomyosin.
DR PRINTS; PRO0194; TROPOMIOSIN.
FT NON TER 502 502
SQ SEQUENCE 502 AA; 55018 MW; 4E73D477CAE79B40 CRC64;

Query Match 96.1%; Score 498; DB 2; Length 502;
Best Local Similarity 97.1%; Pred. No. 1.3e-25;
Matches 101; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Cy 1 LAKQTELEKLDLNDPEGKTODELDKKAABAEIDKKADELPNKVADLEKEISNLEITLG 50
Db 347 LAKQTELEKLDLNDPEGKTODELDKKAABAEIDKKADELPNKVADLEKEISNLEITLG 406

Cy 61 GADPEDDTAALPNKLTATKKALEKTKOKELDALNELGPDGDEE 104
```

```
Db 407 GADPEDDTAALQNKLTATKKALEKTKOKELDALNELGPDGDEE 450

RESULT 4
ID Q9L5B7 PRELIMINARY; PRT; 249 AA.
AC Q9L5B7;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Pspa (Fragment).
GN Name=pspa;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCB1_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=50;
RX MEDLINE=20472698; PubMed=11015380;
RA Beall B., Gherardi G., Facklam R.R., Hollingshead S.K.;
RT "Pneumococcal pspa sequence types of prevalent multiresistant
RT pneumococcal strains in the United States and of internationally
RT disseminated clones.";
RL J. Clin. Microbiol. 38:3663-3669(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=50;
RX Beall B.W.;
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF253405; AAF67353.1; -.
FT NON TER 1 1
SQ SEQUENCE 249 AA; 27271 MW; B4106707EF108A0B CRC64;

Query Match 95.0%; Score 492; DB 2; Length 249;
Best Local Similarity 96.2%; Pred. No. 1.7e-25;
Matches 100; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Cy 1 LAKQTELEKLDLNDPEGKTODELDKKAABAEIDKKADELPNKVADLEKEISNLEITLG 60
Db 103 LAKQTELEKLDLNDPEGKTODELDKKAABAEIDKKADELPNKVADLEKEISNLEITLG 162

Cy 61 GADPEDDTAALPNKLTATKKALEKTKOKELDALNELGPDGDEE 104
Db 163 GADPEDDTAALQNKLTATKKALEKTKOKELDALNELGPDGDEE 206

RESULT 5
ID Q9L585 PRELIMINARY; PRT; 249 AA.
AC Q9L585;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Pspa (Fragment).
GN Name=pspa;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCB1_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=18;
RX MEDLINE=20472698; PubMed=11015380;
RA Beall B., Gherardi G., Facklam R.R., Hollingshead S.K.;
RT "Pneumococcal pspa sequence types of prevalent multiresistant
RT pneumococcal strains in the United States and of internationally
RT disseminated clones.";
RL J. Clin. Microbiol. 38:3663-3669(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=18;
```



RA Beall B.W.;  
 RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF255542; AAF68095.1; -  
 FT NON\_TER 1  
 FT NON\_TER 249  
 SQ SEQUENCE 249 AA; 27050 MW; DFAD2ED9265986FA CRC64;

Query Match 94.8%; Score 491; DB 2; Length 249;  
 Best Local Similarity 96.2%; Pred. No. 2e-25;  
 Matches 100; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 LAKKOTELKLDNDLPBEGKTODELDKKAABELDKKADLPNKVADLEKEISNLEILLG 60  
 DB 68 LAKKOTELKLDNDLPBEGKTODELDKKAABELDKKADLPNKVADLEKEISNLEILLG 127

QY 61 GADPEDDTAALPNKATTKAELEKTOKEIDALNLELGPDDDEE 104  
 DB 128 GADPEDDTAALPNKATTKAELEKTOKEIDALNLELGPDDDEE 171

## RESULT 6

Q9LS90 PRELIMINARY; PRT; 256 AA.

AC Q9LS90;  
 DT 01-OCT-2000 (TRENBLREL. 15, Created)  
 DT 01-OCT-2000 (TRENBLREL. 15, Last sequence update)  
 DT 01-OCT-2003 (TRENBLREL. 25, Last annotation update)  
 DE Pcpa (Fragment).  
 GN Name=pspa;  
 OS Streptococcus pneumoniae.  
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;  
 OC Streptococcus.  
 OK NCBI\_TaxID=1313;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=SP193;  
 RX MEDLINE=20472698; PubMed=11015380;

RA Beall B., Cherardi G., Packiam R.R., Hollingshead S.K.;  
 "Pneumococcal pcpa sequence types of prevalent multiresistant  
 RT pneumococcal strains in the United States and of internationally  
 RT disseminated clones";  
 RT J. Clin. Microbiol. 38:3663-3669 (2000).  
 RL [2]  
 RN SEQUENCE FROM N.A.  
 RC STRAIN=SP193;  
 RA Beall B.W.;  
 RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF254259; AAF68094.1; -  
 FT NON\_TER 1  
 FT NON\_TER 256  
 SQ SEQUENCE 256 AA; 27738 MW; 7F05351559AD9238 CRC64;

Query Match 94.8%; Score 491; DB 2; Length 256;  
 Best Local Similarity 96.2%; Pred. No. 2e-25;  
 Matches 100; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 LAKKOTELKLDNDLPBEGKTODELDKKAABELDKKADLPNKVADLEKEISNLEILLG 60  
 DB 66 LAKKOTELKLDNDLPBEGKTODELDKKAABELDKKADLPNKVADLEKEISNLEILLG 125

QY 61 GADPEDDTAALPNKATTKAELEKTOKEIDALNLELGPDDDEE 104  
 DB 126 GADPEDDTAALPNKATTKAELEKTOKEIDALNLELGPDDDEE 169

## RESULT 7

Q9LS93 PRELIMINARY; PRT; 209 AA.

AC Q9LS93;  
 DT 01-OCT-2000 (TRENBLREL. 15, Created)  
 DT 01-OCT-2000 (TRENBLREL. 15, Last sequence update)  
 DT 01-OCT-2003 (TRENBLREL. 25, Last annotation update)  
 DE Pcpa (Fragment).

GN Name=pspa;

OS Streptococcus pneumoniae.  
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;  
 OC Streptococcus.  
 OK NCBI\_TaxID=1313;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=115;  
 RX MEDLINE=20472698; PubMed=11015380;

RA Beall B., Cherardi G., Packiam R.R., Hollingshead S.K.;  
 "Pneumococcal pcpa sequence types of prevalent multiresistant  
 RT pneumococcal strains in the United States and of internationally  
 RT disseminated clones";  
 RT J. Clin. Microbiol. 38:3663-3669 (2000).  
 RL [2]  
 RN SEQUENCE FROM N.A.  
 RC STRAIN=115;  
 RA Beall B.W.;  
 RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF254256; AAF68091.1; -  
 FT NON\_TER 1  
 FT NON\_TER 209  
 SQ SEQUENCE 209 AA; 22628 MW; 06FF588F7C3BD5B7 CRC64;

Query Match 93.4%; Score 484; DB 2; Length 209;  
 Best Local Similarity 95.2%; Pred. No. 4.9e-25;  
 Matches 99; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 LAKKOTELKLDNDLPBEGKTODELDKKAABELDKKADLPNKVADLEKEISNLEILLG 60  
 DB 19 LAKKOTELKLDNDLPBEGKTODELDKKAABELDKKADLPNKVADLEKEISNLEILLG 78

QY 61 GADPEDDTAALPNKATTKAELEKTOKEIDALNLELGPDDDEE 104  
 DB 79 GADPEDDTAALPNKATTKAELEKTOKEIDALNLELGPDDDEE 122

## RESULT 8

Q9LSB8 PRELIMINARY; PRT; 228 AA.

AC Q9LSB8;  
 DT 01-OCT-2000 (TRENBLREL. 15, Created)  
 DT 01-OCT-2000 (TRENBLREL. 15, Last sequence update)  
 DT 01-MAR-2004 (TRENBLREL. 26, Last annotation update)  
 DE Pcpa (Fragment).  
 GN Name=pspa;  
 OS Streptococcus pneumoniae.  
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;  
 OC Streptococcus.  
 OK NCBI\_TaxID=1313;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=60;  
 RX MEDLINE=20472698; PubMed=11015380;

RA Beall B., Cherardi G., Packiam R.R., Hollingshead S.K.;  
 "Pneumococcal pcpa sequence types of prevalent multiresistant  
 RT pneumococcal strains in the United States and of internationally  
 RT disseminated clones";  
 RT J. Clin. Microbiol. 38:3663-3669 (2000).  
 RL [2]  
 RN SEQUENCE FROM N.A.  
 RC STRAIN=60;  
 RA Beall B.W.;  
 RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF253404; AAF67352.1; -  
 FT NON\_TER 1  
 FT NON\_TER 228  
 SQ SEQUENCE 228 AA; 24430 MW; B6EAA953BC54EA0F CRC64;

Query Match 93.4%; Score 484; DB 2; Length 228;  
 Best Local Similarity 95.2%; Pred. No. 5.3e-25;  
 Matches 99; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

```

QY 1 LAKKQTELEKLDLNDPEGKTQDELDKAEAEALDKKADBLPNKVADLEKEISNLEITLG 60
   |||||
DB 46 LAKKQTELEKLDLNDPEGKTQDELDKAEAEALDKKADBLQNKVADLEKEISNLEITLG 105
   |||||
QY 61 GADPEDDTAALPNKLTATKKALEKTQKELDAALNELGPDGDEE 104
   |||||
DB 106 GADSEDDTAALQNKLTATKKALEKTQKELDAALNELGPDGDEE 149
   |||||

RESULT 9
Q9LS82 PRELIMINARY; PRT; 235 AA.
AC Q9LS82;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE Psppa (Fragment).
GN Name=psppa;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=3;
RX MEDLINE=20472698; PubMed=11015380;
RA Beall B., Gherardi G., Facklam R.R., Hollingshead S.K.;
RT "Pneumococcal psppa sequence types of prevalent multiresistant
RT disseminated clones."
RL J. Clin. Microbiol. 38:3663-3669(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=3;
RA Beall B.W.;
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF255545; AAF68098.1; -.
FT NON_TER 1
FT NON_TER 1
SQ SEQUENCE 235 AA; 25424 MW; BFFBBA8C52CA8380 CRC64;

Query Match 93.4%; Score 484; DB 2; Length 235;
Best Local Similarity 95.2%; Pred. No. 5.5e-25;
Matches 99; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 LAKKQTELEKLDLNDPEGKTQDELDKAEAEALDKKADBLPNKVADLEKEISNLEITLG 60
   |||||
DB 56 LAKKQTELEKLDLNDPEGKTQDELDKAEAEALDKKADBLQNKVADLEKEISNLEITLG 115
   |||||
QY 61 GADPEDDTAALPNKLTATKKALEKTQKELDAALNELGPDGDEE 104
   |||||
DB 116 GADSEDDTAALQNKLTATKKALEKTQKELDAALNELGPDGDEE 159
   |||||

RESULT 10
Q9LS04 PRELIMINARY; PRT; 249 AA.
AC Q9LS04;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE Psppa (Fragment).
GN Name=psppa;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SP195;
RX MEDLINE=20472698; PubMed=11015380;
RA Beall B., Gherardi G., Facklam R.R., Hollingshead S.K.;
RT "Pneumococcal psppa sequence types of prevalent multiresistant

```

```

RT pneumococcal strains in the United States and of internationally
RT disseminated clones."
RL J. Clin. Microbiol. 38:3663-3669(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=SP195;
RA Beall B.W.;
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF252286; AAF69499.1; -.
FT NON_TER 1
FT NON_TER 1
SQ SEQUENCE 249 AA; 26821 MW; F8EA39225CF8DA3F CRC64;

Query Match 93.4%; Score 484; DB 2; Length 249;
Best Local Similarity 95.2%; Pred. No. 5.8e-25;
Matches 99; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 LAKKQTELEKLDLNDPEGKTQDELDKAEAEALDKKADBLPNKVADLEKEISNLEITLG 60
   |||||
DB 67 LAKKQTELEKLDLNDPEGKTQDELDKAEAEALDKKADBLQNKVADLEKEISNLEITLG 126
   |||||
QY 61 GADPEDDTAALPNKLTATKKALEKTQKELDAALNELGPDGDEE 104
   |||||
DB 127 GADSEDDTAALQNKLTATKKALEKTQKELDAALNELGPDGDEE 170
   |||||

RESULT 11
Q9LS83 PRELIMINARY; PRT; 252 AA.
AC Q9LS83;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Psppa (Fragment).
GN Name=psppa;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=127;
RX MEDLINE=20472698; PubMed=11015380;
RA Beall B., Gherardi G., Facklam R.R., Hollingshead S.K.;
RT "Pneumococcal psppa sequence types of prevalent multiresistant
RT pneumococcal strains in the United States and of internationally
RT disseminated clones."
RL J. Clin. Microbiol. 38:3663-3669(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=127;
RA Beall B.W.;
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF255544; AAF68097.1; -.
FT NON_TER 1
FT NON_TER 1
SQ SEQUENCE 252 AA; 27260 MW; 82DE13741F369CA2 CRC64;

Query Match 93.4%; Score 484; DB 2; Length 252;
Best Local Similarity 95.2%; Pred. No. 5.9e-25;
Matches 99; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 LAKKQTELEKLDLNDPEGKTQDELDKAEAEALDKKADBLPNKVADLEKEISNLEITLG 60
   |||||
DB 73 LAKKQTELEKLDLNDPEGKTQDELDKAEAEALDKKADBLQNKVADLEKEISNLEITLG 132
   |||||
QY 61 GADPEDDTAALPNKLTATKKALEKTQKELDAALNELGPDGDEE 104
   |||||
DB 133 GADSEDDTAALQNKLTATKKALEKTQKELDAALNELGPDGDEE 176
   |||||

RESULT 12
Q8X0K3

```

```
ID 08OKK3 PRELIMINARY; PRT; 360 AA.
AC 08OKK3;
DT 01-OCT-2002 (TREMblrel. 22, Created)
DT 01-OCT-2002 (TREMblrel. 22, Last sequence update)
DE 01-MAR-2003 (TREMblrel. 23, Last annotation update)
DB Pneumococcal surface protein A (Fragment).
GN Name=papA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=259/98;
RX MEDLINE=22170754; PubMed=12183557;
RA Miyaji E.N., Ferreira D.M., Lopes A.P.Y., Brandileone M.C.C.,
RA Dias W.O., Leite L.C.C.;
RT "Analysis of serum cross-reactivity and cross-protection elicited by
RT immunization with DNA vaccines against Streptococcus pneumoniae
RT expressing PapA fragments from different clades.";
RL Infect. Immun. 70:5086-5090(2002).
DR EMBL; AF082389; AAL92494.1;
DR InterPro; IPR000533; Tropomyosin.
DR PRINTS; PR00194; TROPOMYOSIN.
FT NON_TER 1
FT NON_TER 360
SQ SEQUENCE 360 AA; 39575 MW; 0C09A791547A7EC CRC64;

Query Match
Best Local Similarity 93.4%; Score 484; DB 2; Length 360;
Matches 99; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 LAKQTELEKLDNDLPBEGKTQDELDEKAEAEALDKKADLPNKVADLEKEISNLEILLG 60
DB 233 LAKQTELEKLDNDLPBEGKTQDELDEKAEAEALDKKADLPNKVADLEKEISNLEILLG 291

QY 61 GADPEDDTAALPNKATATKKALEKTOKEIDALNALNELGPGDDEE 104
DB 292 GADSEDDTAALQNKLATKKALEKTOKEIDALNALNELGPGDDEE 335

RESULT 13
09LAX7 PRELIMINARY; PRT; 429 AA.
AC 09LAX7;
DT 01-OCT-2000 (TREMblrel. 15, Created)
DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)
DE 01-MAR-2004 (TREMblrel. 26, Last annotation update)
DE PapA (Fragment).
GN Name=papA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AC122;
RX MEDLINE=20448953; PubMed=10992499;
RA Hollingshead S.K., Becker R., Briles D.E.;
RT "Diversity of PapA: mosaic genes and evidence for past recombination
RT in Streptococcus pneumoniae.";
RL Infect. Immun. 68:5889-5900(2000).
DR EMBL; AF071818; AAF27714.1;
DR InterPro; IPR011009; Kinase_like.
DR PRINTS; PR00194; TROPOMYOSIN.
FT NON_TER 429
FT NON_TER 429
SQ SEQUENCE 429 AA; 47048 MW; BC1D74BBA54DA9D6 CRC64;

Query Match
Best Local Similarity 93.4%; Score 484; DB 2; Length 429;
Matches 99; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
```

```
QY 1 LAKQTELEKLDNDLPBEGKTQDELDEKAEAEALDKKADLPNKVADLEKEISNLEILLG 60
DB 254 LAKQTELEKLDNDLPBEGKTQDELDEKAEAEALDKKADLPNKVADLEKEISNLEILLG 313

QY 61 GADPEDDTAALPNKATATKKALEKTOKEIDALNALNELGPGDDEE 104
DB 314 GADSEDDTAALQNKLATKKALEKTOKEIDALNALNELGPGDDEE 357

RESULT 14
09LAX9 PRELIMINARY; PRT; 526 AA.
AC 09LAX9;
DT 01-OCT-2000 (TREMblrel. 15, Created)
DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)
DE 01-MAR-2004 (TREMblrel. 26, Last annotation update)
DE PapA (Fragment).
GN Name=papA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=EF3296;
RX MEDLINE=20448953; PubMed=10992499;
RA Hollingshead S.K., Becker R., Briles D.E.;
RT "Diversity of PapA: mosaic genes and evidence for past recombination
RT in Streptococcus pneumoniae.";
RL Infect. Immun. 68:5889-5900(2000).
DR EMBL; AF071816; AAF27712.1;
DR InterPro; IPR011009; Kinase_like.
FT NON_TER 526
FT NON_TER 526
SQ SEQUENCE 526 AA; 58106 MW; 5F1F564A2CB678AB CRC64;

Query Match
Best Local Similarity 93.4%; Score 484; DB 2; Length 526;
Matches 99; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 LAKQTELEKLDNDLPBEGKTQDELDEKAEAEALDKKADLPNKVADLEKEISNLEILLG 60
DB 346 LAKQTELEKLDNDLPBEGKTQDELDEKAEAEALDKKADLPNKVADLEKEISNLEILLG 405

QY 61 GADPEDDTAALPNKATATKKALEKTOKEIDALNALNELGPGDDEE 104
DB 406 GADSEDDTAALQNKLATKKALEKTOKEIDALNALNELGPGDDEE 449

RESULT 15
08VO55 PRELIMINARY; PRT; 608 AA.
AC 08VO55;
DT 01-MAR-2002 (TREMblrel. 20, Created)
DT 01-MAR-2002 (TREMblrel. 20, Last sequence update)
DE 01-MAR-2004 (TREMblrel. 26, Last annotation update)
DE Pneumococcal surface protein A (Fragment).
GN Name=papA;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=KNH1156;
RX Lee K.J., Bae S.M., Chung K.S.;
RT Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
RL EMBL; AF460993; AAL67804.1;
DR InterPro; IPR002479; CW_binding.
DR PRINTS; PR000566; Lipocin_cytfabp.
DR Pfam; PF01473; CW_binding_1; 10.
DR PROSITE; PS00213; LIPOCALIN; UNKNOWN_2.
FT NON_TER 1
FT NON_TER 608
```

SQ SEQUENCE 608 AA; 67918 MW; 15F71BD62E297526 CRC64;

Query Match	93.4%	Score 484;	DB 2;	Length 608;
Best Local Similarity	95.3%	Pred No 1	48-24;	

Matches 99; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Matches 99; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

```

09      1 LAKKOTELEKLDNDLDPBEGKQODELKEAAREALDKADELNNKVALEKEISNLEITLIG 60
Dbb    222 LAKKOTELEKLDSDLPBEGKQODELKEAREAEALDKADELQNKVALEKEISNLEITLIG 2811

```

Db 222 LAKKOTELEKLLDSDPEGKTQDELDKEAEAEALDKADELQNKVADLEKEISNLEILG 281

61 GADPEDDTAALPNKLTATKKAELEKTQKEIDALNELGPDGDEEE 104

Db 282 GADSEDDTAALQNKLATKKAELEKTQKELDALNELGPDGDEEE 325

Search completed: November 19, 2004, 22:31:53  
Job time : 72.8211 secs

Job time : 72.8211 secs